非线性科学丛书

分形与图象压缩

陈守吉 张立明 编著

上海科技教育出版社

本书出版由上海市新闻出版局
学术著作出版基金资助

非线性科学丛书

分形与图象压缩

陈守吉　张立明　编著

上海科技教育出版社

内 容 提 要

本书是非线性科学丛书中的一种，介绍分形在图象压缩编码中的应用. 全书计分三章，包括分形几何基础、迭代函数系统、拼贴定理、分形图象压缩的基本原理和实现方法，本书是一本非线性科学应用于图象压缩的科技著作.

本书可供理工科大学教师、高年级学生、研究生阅读，也可供有关研究人员参考.

非线性科学丛书

分形与图象压缩

陈守吉　张立明 编著

上海科技教育出版社出版发行

（上海市冠生园路393号　邮政编码：200233）

各地新华书店经销　　上海市印刷十二厂印刷

开本：850×1168　1/32　　印张：4.625　字数：112 000

1998年12月第1版　　1998年12月第1次印刷

印数：1—3200

ISBN 7-5428-1642-X/O・164　　定价：（精装本）11.50元

非线性科学丛书编辑委员会

主　编：郝柏林

副主编：郑伟谋　吴智仁

编　委：（按姓氏笔画为序）

丁鄂江　文志英　朱照宣

刘式达　刘寄星　孙义燧

杨清建　李邦河　张洪钧

张景中　陈式刚　周作领

赵凯华　胡　岗　顾　雁

倪皖荪　徐京华　郭柏灵

陶瑞宝　谢惠民　蒲富恪

霍裕平　魏荣爵

非线性科学丛书

出版说明

现代自然科学和技术的发展，正在改变着传统的学科划分和科学研究的方法．“数、理、化、天、地、生”这些曾经以纵向发展为主的基础学科，与日新月异的新技术相结合，使用数值、解析和图形并举的计算机方法，推出了横跨多种学科门类的新兴领域．这种发展的一个重要特征，可以概括为“非”字当头，即出现了以“非”字起首而命名的一系列新方向和新领域．其中，非线性科学占有极其重要的位置．这决非人们“想入非非”，而是反映了人类对自然界认识过程的螺旋式上升．

曾几何时，非线性还被人们当作个性极强，无从逾越的难题．每一个具体问题似乎都要求发明特殊的算法，运用新颖的技巧．诚然，力学和数学早就知道一批可以精确求解的非线性方程，物理学也曾经严格地解决过少数非平庸的模型．不过，这些都曾是稀如凤毛麟角的“手工艺”珍品，人们还没有悟出它们的普遍启示，也没有看到它们之间的内在联系．

20 世纪 60 年代中期，事情从非线性现象的两个极端同时发生变化．一方面，描述浅水波运动的一个偏微分方程的数值计算，揭示了方程的解具有出奇的稳定和保守性质．这启发人们找到了求解一大类非线性偏微分方程的普遍途径，即所谓“反散射”方法．反散射方法大为扩展了哈密顿力学中原有的可积性概念，反映了这类方程内秉的对称和保守性质．到了 80 年代，反散射方法推广到量子问题，发现了可积问题与统计物理中严格可解模型的联系．

60年代初期还证明了关于弱不可积保守系统普遍性质的KAM定理. 于是,非线性问题的可积的极端便清楚勾划出来,成为一个广泛的研究领域. 虽然这里的大多数进展还只限于时空维数较低的系统,但它对非线性科学发展的促进作用是不可估量的.

另一方面,在"不可积"的极端,对KAM定理条件的"反面文章",揭示了保守力学系统中随机性运动的普遍性,而在耗散系统中则发现了一批奇怪吸引子和混沌运动的实例. 这些研究迅速地融成一片,一些早年被认为是病态的特例也在新的观点下重新认识. 原来不含有任何外来随机因素的完全确定论的数学模型或物理系统,其长时间行为可能对初值的细微变化十分敏感,同投掷骰子一样地随机和不可预测. 然而,混沌不是无序,它可能包含着丰富的内部结构.

同时,由于计算科学特别是图形技术的长足进步,人们得以理解和模拟出许多过去无从下手研究的复杂现象. 从随机与结构共存的湍流图象,到自然界中各种图样花纹的选择与生长,以及生物形态的发生过程,都开始展现出其内在的规律. 如果说,混沌现象主要是非线性系统的时间演化行为,则这些复杂系统要研究的是非线性地耦合到一起的大量单元或子系统的空间组织或时空过程. 标度变换下的不变性、分形几何学和重正化群技术在这里起着重要作用.

在由上述种种方面汇成的非线性科学洪流中,许多非线性数学中早已成熟的概念和方法开始向其他学科扩散,同时也提出了新的深刻的数学问题. 物理学中关于对称和守恒,对称破缺,相变和重正化群的思想,也在日益增多的新领域中找到应用. "非线性"一词曾经是数学中用以区别于"线性"问题的术语,非线性科学正在成为跨学科的研究前沿. 各门传统学科中都有自己的非线性篇章,非线性科学却不是这些篇章的总和. 非线性科学揭示各种非线性现象的共性,发展处理它们的普适方法.

这样迅猛发展的跨学科领域,很难设想用少数专著加以概括,

何况学科发展的不少方面还未成熟到足以总结成书的地步. 于是,有了动员在前沿工作的教学和研究人员,以集体力量撰写一套"非线性科学丛书"的想法. 在上海科技教育出版社的大力支持下,这一计划得以付诸实现.

这套"非线性科学丛书"不是高级科普,也不是大块专著. 它将致力于反映非线性科学各个方面的基本内容和最新进展,帮助大学高年级学生、研究生、博士后人员和青年教师迅速进入这一跨学科的新领域,同时为传统自然科学和工程技术领域中的研究和教学人员更新知识提供自学教材. 非线性科学的全貌将由整套丛书刻划,每册努力讲清一个主题,一个侧面,而不求面面俱到,以免失之过泛. 在写作风格上,作者们将努力深入浅出;图文并茂,文献丰富;力求有实质内容,无空洞议论,以真刀真枪脚踏实地武装读者. 从读者方面,自然要求具备理工科大学本科的数学基础,和读书时自己主动思索与推导的习惯.

"非线性科学丛书"的成功,取决于读者和作者的支持. 我们衷心欢迎批评和建议.

郝柏林

1992年4月30日于北京中关村

前言

分形是非线性科学中的一个重要领域. 分形图象是一种具有复杂几何形状,不规则的图象,但其内部存在着无穷多个自相似性,因而可以用一组简单的迭代函数方程通过随机迭代而得到. 这个思想在 80 年代末被引入到图象的压缩编码中,从严格具有自相似性的分形图象推广到一般的任意图象,从黑白图象推广到有灰阶甚至于彩色图象上. 如果任意图象都可以近似为分形图象,那么只要找到其图象内部存在的自相似迭代函数的参数,则图象就可用迭代函数的参数来表达,这就大大压缩了图象的信息量,解决图象压缩编码中的问题. 近年来,由于信息科学和工程应用的发展,图象分形编码的研究也发展很快,分形几何在图象压缩缩码中的应用成为十分诱人的研究领域.

本书旨至介绍分形几何学在图象编码应用中的基本理论,基本方法和最新进展. 本书不是全面地介绍分形理论,而且针对有关信息压缩编码尤其是数字图象编码的应用展开的. 本书中的不少内容是我们研究小组包括一些年青的研究生几年的研究计算机模拟的结果,我们希望介绍给读者,能使读者了解非线性科学在工程应用方面的进展,以及让从事信息和图象编码的工程技术人员了解非线性的科学方法,并能迅速跨入这个新领域.

本书共分为三章. 第 1 章主要介绍分形几何及分形用于图象压缩编码的数学基础和理论依据. 第 2 章主要介绍分形图象编码的基本方法及最新进展. 第 3 章介绍分形编码和其他方法相结合的进展,如小波变换与分形的结合方法,余弦变换与分形结合的方法等.

本书中不少内容是我们研究小组的顾炜老师、研究生施一蓬、王炜、李莉等研究和模拟的结果. 本书还得到复旦大学非线性中心的同仁们的支持，至此一并表示谢意.

由于我们学识肤浅，不妥和错误之处在所难免，望读者批评指正. 目前分形图象编码还在研究之中，随着时间发展一定会有更好的理论和方法涌现，希望本书起一个抛砖引玉的作用.

陈守吉、张立明

1998 年 12 月

Abstract

A survey in Image Compression Coding by using Fractal, main area of nonlinear science, is presented in this book, which consists of three chapters (eleven sections).

As nonlinear science book applied in Image compression, this book includes Fractal Geometry, Iterated Function System, Collage theorem, principle and methods of Fractal image compression etc.

Readership includes teachers in college or university, Senior undergraduate students and graduate students. It is also as reference for some researchers and engineers.

目　　录

Contents

第1章

分形几何和图象压缩基础

§1 欧几里得几何学与分形几何学

众所周知，欧几里得几何学研究的图形都是规则的形状，例如圆、正方形、球和圆锥体等等，构成这些图形的边缘（线或者面），都是连续而光滑的．根据欧几里得几何学，可以方便地描述砖块、轮子、机器部件以及建筑物等等．但是，大自然中的许多形状都很不规则，甚至是支离破碎的．例如天空中的云彩不是球体，地面上的海岸线不是圆弧，山脉不是锥体，树皮不是光滑的曲面，动物体内血管的分布更是错综复杂．这些不规则的几何形状也经常出现在自然科学的各个领域中．例如流体力学中的湍流，物理学中的布朗（R. Brown）运动，化学中酶的构造，生物学中细胞的生长，非线性动力学中的奇怪吸引子以及工程技术中的信号处理等等．为了研究这些大自然的几何学，就诞生了一门新的分支——分形几何学．

分形几何学是由曼德勃罗特（B. B. Mandelbrot）在20世纪70年代创立的．"分形（fractal）"一词，也是由曼德勃罗特提出的．它来源于拉丁语"fractus"，含有"不规则"和"破碎"的意义．

描述规则形状的欧几里得几何学提供了构造物理对象的一级近似，是用于传递技术产品构思的语言；而分形几何学则为研究一类非规则几何对象提供了思想、方法和技巧等方面的总框架，它可以构造从植物到星系的物理结构的精确模型，它是一种新的语言．曼德勃罗特曾指出，"分形语言和'老的'欧几里得语言分别为完全不同的目标服务．"[2]

分形理论（包括分形几何学、分形物理学等）是非线性科学的

主要分支之一.它在自然科学各个学科中,甚至在经济和社会活动中,都有着广泛的应用.本书所介绍的利用分形理论进行图象压缩,就有着重要的实用意义.

大自然中的所有形状和人们考虑的一切图形(可统称为几何体),可以分为两大类.一类是具有特征尺度的,例如人的身高,球的半径,建筑物的长、宽、高,等等.具有特征尺度的几何体有一个重要性质,即构成几何体的线或面都是光滑的.另一类是没有特征尺度的,即必须同时考虑从小到大的许许多多尺度.例如夏季天空中翻滚的积雨云,北方冬季玻璃窗上的冰霜,以及极为普遍的湍流现象(小至静室中缭绕的青烟、大到木星大气中的涡流).这些所谓"无标度"的几何体,其实就是分形.它们也有一个共同的特征,即自相似性.

曼德勃罗特曾经指出[2],[4],分形具有三个要素:形状(form),机遇(chance)和维数(dimension).首先,前已指出,分形的形状是支离破碎、参差不齐和凹凸不平的不规则形状;其次,我们发现大自然中的海岸线与用以描述它的科克分形曲线(见本书第7页例2)之间仍有很大不同,而这种差异是由于海岸线受到自然界随机因素的作用而产生的.同时,巴恩斯利(M.F. Barnsley)发现,可以对一组给定的规则通过随机迭代而得到分形(见§2),而对象本身并不依赖于随机性,我们总是以百分之百的概率得到同一个分形,因此随机性或者机遇仅仅是工具,而结果却是确定性的;第三,分形的维数可以是分数.这是一种新的维数,称为分维.维数是几何对象的一个重要特征量.通常的"维数"概念,指的是为了确定几何对象中一个点的位置所需要的独立坐标的数目.在这种意义下,它是一个整数.所以点是0维的,线是1维的,面是2维的,而立方体是3维的,等等.由于这种定义具有对几何对象在同胚变换下的不变性,因此称为"拓扑维",记作 D_T.维数研究的突破点,是曼德勃罗特探索了把一类集合(其维数可以是分数)应用到大自然科学现象(从分子到星系)的模型中去[2],[3],[6].

定义 1[4] 所有离散集合的**拓扑维** D_T 都为0. 如果一个集合 F 中的每一点的任意小邻域的边界都具有拓扑维 $D_T = n-1$，这里 n 是整数，那么这个集合 F 的**拓扑维** D_T 是 n.

拓扑维的特点是直观性强，与人们的经验相符合. 但对有些几何对象(例如分形)会产生矛盾. 早在1919年，豪斯多夫(F. Hausdorff)提出了维数不必限于整数，而可以是个实数，称为**豪斯多夫维**(又称为**豪斯多夫-伯西柯维奇**(A. S. Besicovitch)**维**)，记作 D. 为此，先介绍集合的"覆盖"和"豪斯多夫测度"的概念.

设 U 是 n 维欧氏空间 $\mathbb{R}^n$ 中的非空子集. 我们把 U 内任何两点间距离的最大值，定义为 U 的**直径**，记作 $|U|$. 即

$$|U| = \sup\{|x-y| : x, y \in U\}.$$

对于集 $F \subset \mathbb{R}^n$，如果 F 可以被可列个(或有限个)直径不超过 δ 的集类 $\{U_i\}$ 所覆盖，即

$$F \subset \bigcup_i U_i, \quad 且 \quad 0 < |U_i| \le \delta, \qquad \forall i$$

则称 $\{U_i\}$ 是 F 的一个 δ-覆盖.

设 $F \subset \mathbb{R}^n$，s 是一非负实数. 对任何 $\varepsilon > 0$，定义

$$H^s_\varepsilon(F) = \inf\left\{\sum_{i=1}^{\infty} |U_i|^s : \{U_i\} 为 F 的 \delta\text{-覆盖}\right\}, \qquad (1.1)$$

其中要取遍所有直径不超过 ε 的 F 的覆盖而得下确界，且 H^s_ε 当 ε 减小时增加. 关于 ε 取极限，称

$$H^s(F) = \lim_{\varepsilon \to 0} H^s_\varepsilon(F) \qquad (1.2)$$

为 F 的 **s 维豪斯多夫测度**. 对任何 s，上述极限都是存在的. 但通常是0或 ∞.

豪斯多夫测度推广了长度、面积和体积等概念. 可以证明[5]，$\mathbb{R}^n$ 中任何子集的 n 维豪斯多夫测度与通常的 n 维体积，只相差一个常数倍(当 n 是整数时).

定义 2[5] 集 F 的**豪斯多夫维**为

$$D = \inf\{s : H^s(F) = 0\} = \sup\{s : H^s(F) = \infty\}. \qquad (1.3)$$

所以，

$$H^s(F)=\begin{cases}\infty, & \text{当 } s<D \text{ 时};\\ 0, & \text{当 } s>D \text{ 时};\\ [0,\infty) \text{ 或 } \infty, & \text{当 } s=D \text{ 时}.\end{cases} \tag{1.4}$$

如把 F 的豪斯多夫测度 $H^s(F)$ 作为 s 的函数，作出其图象（图 1-1），可以看出，F 的豪斯多夫维数 D 是使得 $H^s(F)$ 从 ∞ 跳跃到 0 的一个临界点. 显然，D 是非负实数，它不一定是整数，所以是一种分维.

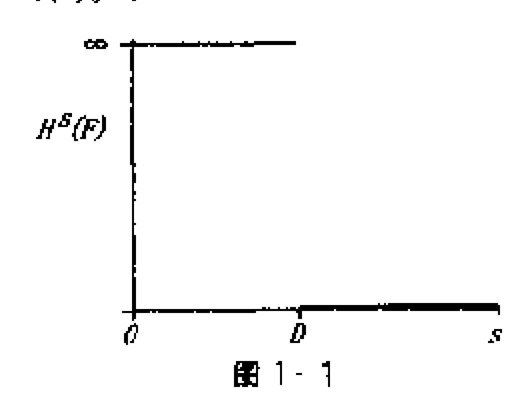


图 1-1

豪斯多夫维具有对任何集 F 都有定义的优点，而缺点主要是难以计算或估计它的值. 但是，它可能是多种多样的分维定义中最重要的一种，也是分维定义中最常用的一种. 本书中，如不特别说明，分维即指豪斯多夫维.

对于欧氏空间 $\mathbb{R}^n$ 中的 F 集，两种维数 D_T 和 D 的相同处是其值都是至少为 0 而至多为 n；但维数 D_T 总是一个整数，而 D 却不必是整数. 两种维数也无需相等，只要满足苏比尔拉（E. Szpilrajn）不等式[4]

$$D \geqslant D_T. \tag{1.5}$$

即任一集合（几何体）的豪斯多夫维总不小于拓扑维.

下面再介绍两种维数：盒维和相似维.

定义 3[5] 设集 $F \subset \mathbb{R}^n$. 记 $N_\varepsilon(F)$ 是可以覆盖 F 的、边长为 ε 的 n 维立方体（记作 ε^n-立方体）的最少个数，则 F 的**盒维** D_B 定义为（当极限存在时）：

$$D_B=\lim_{\varepsilon\to 0}\frac{\log N_\varepsilon(F)}{-\log \varepsilon}. \tag{1.6}$$

所以，盒维与豪斯多夫维一样，也是考虑 F 的覆盖，只是可采用同样大小的 ε^n-立方体. 所以，总成立[4]

$$D_B \geqslant D. \tag{1.7}$$

对于许多"相当规则"的集，有 $D_B = D$；但也有大量使（1.7）中不

等号严格成立的例子.

盒维的实际计算和估计都比较容易,因此它是应用最广泛的一种维数.

如果集 F 具有"自相似"的子集(即这些子集有着与集 F 相同的结构,而且其大小是按比例缩小的,缩小因子为 r),而这些子集又依次具有相同比例的子子集,如此等等.假定每一个"母集" F 有 N 个子集,而当每一个子集中各点间的距离按因子 $\left(\frac{1}{r}\right)$ 增大时都恒等于该"母集".如果规定初始尺度为 1,并取边长 $\epsilon = r^k$,则覆盖 F 所需立方体的个数 $N_\epsilon(F) = N^k$.代入(1.6)式,有

$$D_B = \lim_{k \to \infty} \frac{\log(N^k)}{-\log(r^k)}.$$

即得到十分简单的维数公式(甚至不需计算极限):

$$D_B = \frac{\log N}{\log \frac{1}{r}}. \tag{1.8}$$

其中 N 是由前一个"母集"按 r 比例缩小后的子集个数,而且取 $D_B = D$.我们把这种集 F 称为**严格的自相似集**.

下面计算一些典型的集合的各种维数.

[例 1] 康托(G. Cantor)集.

可以这样来构造康托集:将单位线段[0,1]三等分,然后挖去中间三分之一,即 $\left(\frac{1}{3}, \frac{2}{3}\right)$,注意,挖掉的是开区间;然后,把剩下的每段再作三等分并挖去各自的中间三分之一;……如此无限次地等分和挖去(图 1-2),最后的极限集就是一种**康托集**(常称为**三分康托集**).它是一维欧氏空间 中无穷多个点的集合.由此就不难算出康托集的各种维数.

1. 康托集是由离散点构成的,可见康托集的拓扑维 $D_t = 0$.

2. 可以算出,康托集的豪斯多夫维

$$D = \log 2 / \log 3 = 0.630929\cdots\cdots.$$

关于它的严格计算,可见[5].

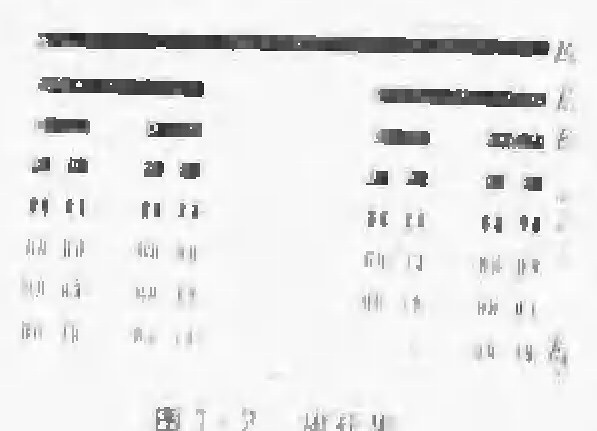


图 1-2　康托集

3. 因为康托集是严格的自相似集，每个"母集"有 2 个子集，均按三分之一比例缩小，所以 $N=2$，$r=\frac{1}{3}$。由式(1.8)即得康托集的盒维 $D_B=\log 2/\log 3$。

由以上可见，对于康托集，

$$D>D_T，且\ D=D_B.$$ □

现在回到"分形"的概念。目前还没有一致公认的"分形"的定义。通常，把具有下面典型性质的集 F 称为**分形**[1]：

1. F 具有精细结构，即它包含有任意小比例的细节；

2. F 具有某种自相似的结构，它可能是精确的，或是近似的，也可能是统计意义下的自相似；

3. F 的分维(以某种方式定义)严格大于它的拓扑维；

4. 存在对 F 的简单算法(低信息量)描述，例如可以由迭代产生。

上列四条性质中，最重要的是第 3 条。曼德勃罗特曾单独把这条作为分形的定义[2]。

下面讨论一些典型的分形。为了便于计算其分维，我们的讨论限于严格的自相似集。这些例子在后面讨论图象压缩时还会遇到。



图 1-3　土星光环

显然，例 1 的康托集就是分形，因为它满足上述的分形特性。由于它就像是散布在直线段上的一些灰尘，曼德勃罗特把它称为**康托尘埃**(Cantor dust)[2]。大自然中的许多现象可用康托尘埃来描述。例如，美丽的土星光环由无限多个细光环构成，它们的分布(在光环

的横截面上)就是某种康托尘埃(见图 1－3).

[例 2] 科克(H. V. Koch)曲线.

设 E_0 是单位长度的线段. 把 E_0 三等分,以其中间三分之一为底边作向外凸出的正三角形,然后去掉该底边即得 E_1;再把同样的过程应用到 E_1 的每个线段(包含四个线段),就得出 E_2;以此类推,直至无穷(图 1－4),其极限曲线 F 就是**科克曲线**. 因为它很像雪花的外形,所以又称为**雪花曲线**. 显然,它也是严格自相似集. 由作图过程看出,$N=4$, $r=\frac{1}{3}$. 由(1.8)式,得

$$D=D_s=\frac{\log 4}{\log 3}=1.261858\cdots\cdots.$$

而其拓扑维 $D_t=1$. 现在 $D>D_t$,所以是分形曲线. 大自然中弯弯曲曲的海岸线,形状很复杂,难以用普通的光滑曲线来描述,因为把任意一段海岸线放大,又含有许许多多凸出的海角和凹陷的海湾. 因此,不能用长度来描述海岸线的复杂性(长度将趋于无限),而雪花曲线可以作为海岸线的很好的模型. 前已算得雪花曲线的分维 $D\approx1.26$,根据测量,实际海岸线的分维也在 1.2 附近,所以分维正是海岸线复杂性的一种度量. □

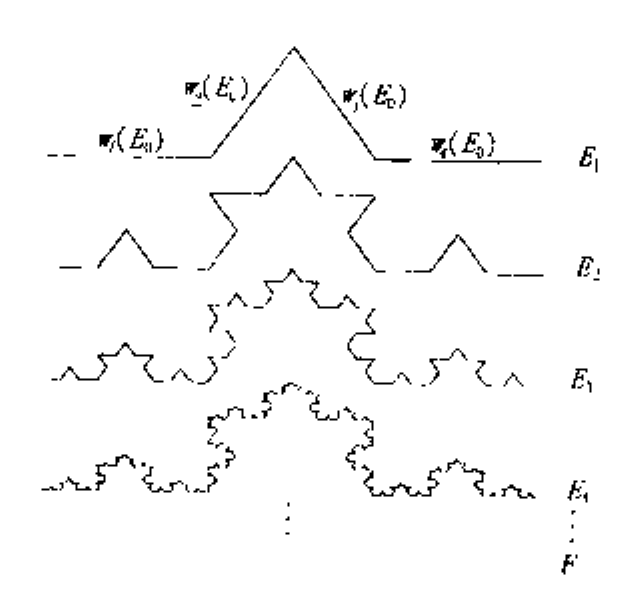
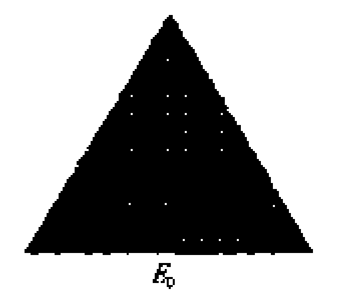
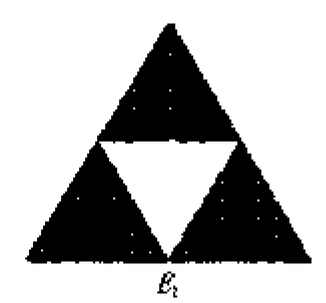


图 1－4　科克曲线

[例 3] **谢尔宾斯基**(W. Sierpinski)**垫片**. 又称为**谢尔宾斯基三角形**.

设 E_0 是 $\mathbb{R}^2$ 中的一个正三角形,把它的三边中点相连,就成为四个小三角形,去掉中间一个倒三角形,就得到 E_1;再把同样的过程应用到 E_1 的每个小三角形,就得到 E_2;以此类推(图 1－5),就得到极限集 F,它像是打有许许多多大大小小三角形孔的垫片,简直是"体无完肤"(没有整块面积),它的拓扑维 $D_t=1$. 显然,它也


E_0


E_1

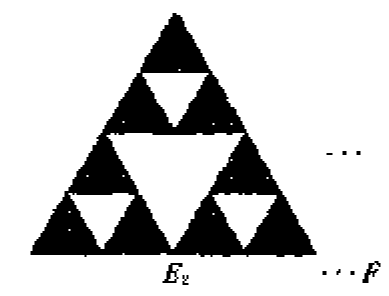

E_2 …F

图 1-5 谢尔宾斯基垫片

是严格自相似的分形，且有 $N=3$，$r=\frac{1}{2}$，所以

$$D = D_B = \frac{\log 3}{\log 2} = 1.584964\cdots\cdots > D_T. \qquad \square$$

［例 4］皮亚诺(G. Peano)曲线.

这是一条可以填满一个正方形的分形曲线. 图 1-6 表示了构造皮亚诺曲线的头几步，图上虚点构成的正方形代表将被皮亚诺曲线填满的面积. 首先把该正方形等分为 4 个，用 3 条线段把这 4 个小正方形的中心相连(设每条线段的长度为 1)，构成曲线 H_0. 现在把 H_0 缩小 1/2 比例后复制 4 份分别置于 4 个小正方形中，并把第 1 个复制品顺时针旋转 90°，而把第 4 个复制品逆时针旋转 90°，然后我们用 3 条小线段(长度为 1/2)如图示那样连接这 4 条小曲线(复制品)的起点和终点，而这最后的曲线就称为 H_1. 第 2 步，我们把 H_1 缩小 1/2 比例后复制 4 份，像第 1 步那样放置到 4 个小正方形中，再用 3 条线段(长度为 1/4)相连，就得到曲线 H_2. H_2 含有 16 个 H_0 的复制品，每个已缩小 1/4. 如此不断继续，第 n 步是把第 $n-1$ 步得到的 H_{n-1} 缩小 1/2 比例后复制 4 份放到规定的位置，并用 3 条长为 $1/2^n$ 的线段连接成整条曲线，H_n 由 4^n 个(按 $1/2^n$ 比例缩小的)H_0 的复制品构成. 当 n 趋于无穷时，它的极限曲线就是充满整个虚点正方形的皮亚诺曲线.

根据以上构造方法，皮亚诺曲线是严格自相似的，即 $N=4$，$r=\frac{1}{2}$，所以，它的分维为

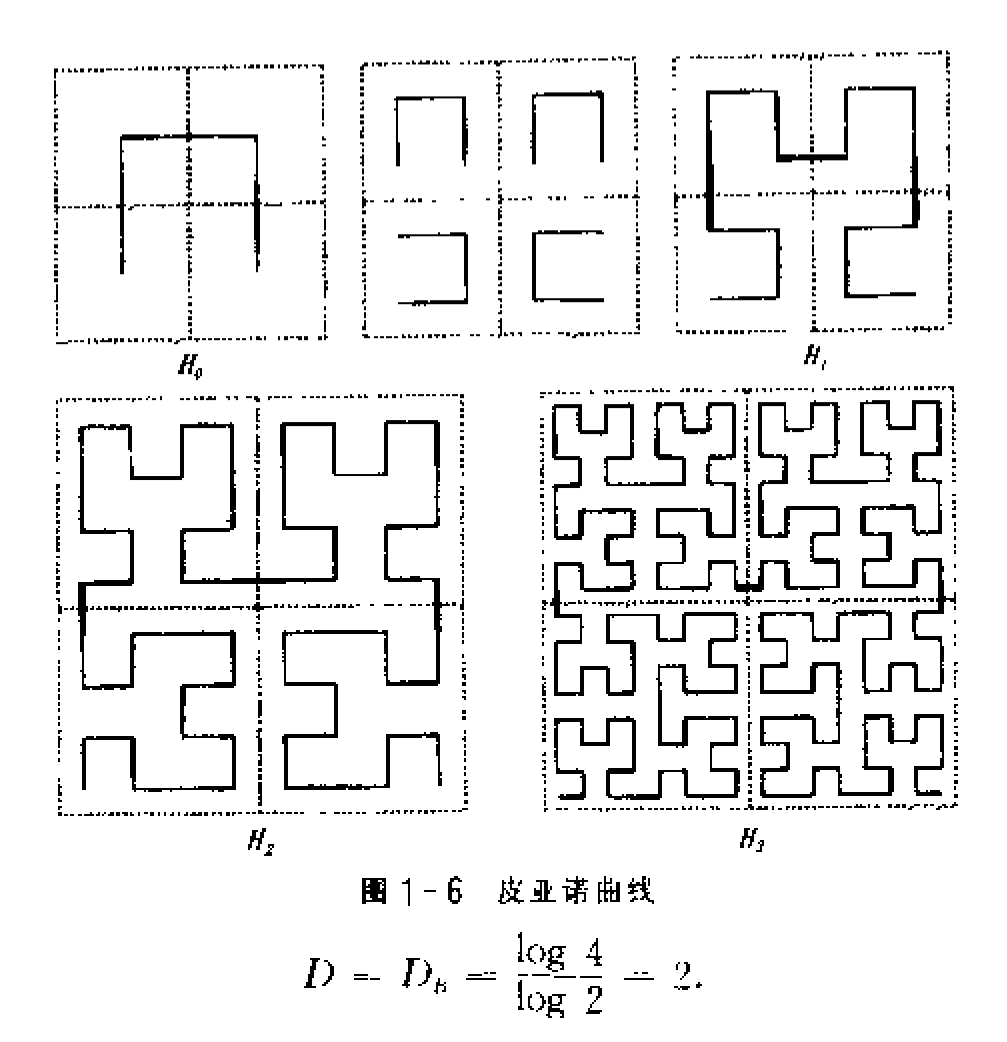


图 1－6　皮亚诺曲线

$$D = D_b = \frac{\log 4}{\log 2} = 2.$$

但它的拓扑维 $D_t = 1$，有 $D > D_t$．因此，皮亚诺曲线是分形，而且其分维是整数.

由于皮亚诺曲线具有填满平面的性质，它在二维的图象压缩技术中有用.

§2　迭代函数系统

现在我们转到讨论一些特殊的分形．它们可以由计算机图形学和数学构造两方面提出．其中最简单的一类，称为自相似集．起初，曼德勃罗特在 1977 年用初始和标准折线的迭代过程来构造这类集[4]．到 1981 年，哈钦森(J. E. Hutchinson)[illegible]推广此迭代过程，首先引入了迭代函数系统(Iterated Function Systems，简称 IFS)的

理论（术语 IFS 是由巴恩斯利于 1985 年提出的[2]），把度量空间中的压缩变换集作为动力系统的模型. 巴恩斯利又把 IFS 的思想应用到图象压缩中. IFS 在分形图象压缩的发展中起着引导的作用，分形图象压缩的大多数工作都是基于 IFS 或它的推广.

一个分形集一般都由无限多个点组成，它们的分布又是如此之复杂，以至不可能通过直接给定每个点的位置来描述它. 因此，就想通过"各部分之间的相互关系"来定义它（例如例 1 中介绍的康托集，就可以看成是由与它自身相似的两部分组成的）. 而迭代函数系统可达到这样的目标.

为了定义迭代函数系统，先要介绍"完备度量空间"和"压缩映射"等基本概念.

1. 分形空间

为了简单起见，只讨论 ^{2} 中的分形，并把研究分形几何的基本空间记作（ ，d ）. 其中 $\subset$ ^{2}，是非空的闭集；d 是 中的度量，如可取 d 为欧氏距离，即

$$d(x,y) = |x - y|, \qquad x,y \in$$

所以（ ，d ）是度量空间. 一般还要求基本空间是完备的. 由于我们讨论的"黑-白"图形或图画都是该空间 的子集，自然还要定义新的空间 $\mathcal{H}$.

设（ ，d ）是完备度量空间，把 中非空紧子集（即有界闭集）的全体记作 $\mathcal{H}$（ ）. 我们可以把 $\mathcal{H}$（ ）想象成 中全体黑-白图形的集合，因此，任意一幅图形都可由 $\mathcal{H}$（ ）的一个子集来表示，该子集所有的点都是黑色的，其他地方是白色的. 所以空间 $\mathcal{H}$ 包含有数量十分巨大的元素.

现在要定义 $\mathcal{H}$（ ）上的度量，即任意两幅图形（子集）之间的"距离"（称为豪斯多夫距离，记作 h ）. 它依赖于基本空间 的度量 d .

定义 1 两个子集间的**豪斯多夫度量** h 定义为

$$h(A,B) = \max\{\sup_{x\in A} d(x,B), \sup_{y\in B} d(y,A)\},$$

$$A,B \in \mathcal{H} \qquad (2.1)$$

其中，从点 x 到子集 B 之间的距离定义为

$$d(x,B)=\inf\{d(x,y);\ y\in B\},$$

$$\forall\ x\in A,B\in\mathscr{H} \qquad (2.2)$$

由于 $\mathscr{H}(X)$ 的子集之间定义了度量 h，$(\mathscr{H}(X),h)$ 也是一个度量空间. 有些文献中把 $(\mathscr{H}(X),h)$ 称为分形空间，因为平面分形都是该空间的元素.

定理 1 设 (X,d) 是完备的度量空间，则 $(\mathscr{H}(X),h)$ 也是完备的度量空间. 而且，如 $\{A_n\in\mathscr{H}(X);\ n=1,2,\cdots,\infty\}$ 是柯西(A. L. Cauchy)序列，则

$$A=\lim_{n\to\infty}A_n\in\mathscr{H}(X)$$

可如下法求得：

$$A=\{x\in X:\ 存在一个收敛于 x 的柯西序列 \{x_n\in A_n\}\}.$$

定理 1 的证明可见[9].

2. 分形空间上的压缩映射

定义 2 设 $w: X\to X$ 是基本空间 (X,d) 上的一个映射. 如果存在一个正的常数 $c<1$，使

$$d(w(x),w(y))\leqslant c\cdot d(x,y),$$

$$x,y\in X \qquad (2.3)$$

则称 w 为 (X,d) 上的**压缩映射**，c 称为**压缩因子**.

图 2-1 表示一张笑脸 F 经压缩映射 w 后变为较小的笑脸 $w(F)$，可以发现，变换后的 $w(F)$ 上的两眼比 F 上的更为靠近了. 实际上，压缩映射总把 F 上的一对点移到更为靠近的一对点.



图 2-1 压缩映射 w

由此就可构造出分形空间 $(\mathscr{H}(X),h)$ 上的压缩映射.

引理 1 设 $w: X \to X$ 是基本空间 (X, d) 的一个压缩映射，带有压缩因子 c. 则由下式

$$\hat{w}(B) = \{w(x): x \in B\} \qquad \forall B \in \mathscr{H}(X) \quad (2.4)$$

定义的映射 $\hat{w}: \mathscr{H}(X) \to \mathscr{H}(X)$，是分形空间 $(\mathscr{H}(X), h)$ 上的压缩映射，且压缩因子也是 c，即

$$h(\hat{w}(A), \hat{w}(B)) \leqslant c \cdot h(A, B). \qquad \forall A, B \in \mathscr{H}(X) \quad (2.5)$$

引理 2 考虑 $(\mathscr{H}(X), h)$ 上 n 个压缩映射 $\{w_i: \mathscr{H}(X) \to \mathscr{H}(X), i=1,2,\cdots,n\}$，其压缩因子分别为 $c_i (i=1,2,\cdots,n)$. 现定义一个新的映射 $W: \mathscr{H}(X) \to \mathscr{H}(X)$，即

$$W(B) = \hat{w}_1(B) \cup \hat{w}_2(B) \cup \cdots \cup \hat{w}_n(B) = \bigcup_{i=1}^{n} \hat{w}_i(B), \qquad \forall B \in \mathscr{H}(X) \quad (2.6)$$

则 W 是压缩映射，且压缩因子 $c = \max\{c_1, c_2, \cdots, c_n\}$. 即

$$h(W(A), W(B)) \leqslant c \cdot h(A, B). \qquad \forall A, B \in \mathscr{H}(X) \quad (2.7)$$

证明 只要对 $n=2$ 的情形证明就可以了.

我们以 A_ε 表示集 A 的 ε-邻域，即表示与集 A 的距离不大于 ε 的点的集合，表为 $A_\varepsilon = \{x$: 对于 A 中的某点 $y, d(x,y) \leqslant \varepsilon\}$.

现选定 $A, B \in \mathscr{H}(X)$，并令 $\varepsilon = h(A, B)$，则由豪斯多夫距离的定义，可知 $A \subset B(\varepsilon)$ 和 $B \subset A(\varepsilon)$，又设 $c = \max(c_1, c_2)$，$A_i = w_i(A)$，$B_i = w_i(B)$ (其中 $i = 1, 2$)，$A' = W(A) = A_1 \cup A_2$ 以及 $B' = W(B) = B_1 \cup B_2$. 现在要证明 $h(W(A), W(B)) \leqslant c \cdot h(A, B)$，也就是要证明 $B' \subset A'(c\varepsilon)$ 和 $A' \subset B'(c\varepsilon)$. 这只要从下列各式就可推出：$B_i \subset w_i(A(\varepsilon)) \subset A_i(c_i\varepsilon) \subset A_i(c\varepsilon)$ 以及 $A_i \subset w_i(B(\varepsilon)) \subset B_i(c_i\varepsilon) \subset B_i(c\varepsilon)$，其中 $i = 1, 2$. □

3. 迭代函数系统

定义 3 完备的度量空间 (X, d) 以及 n 个压缩映射 $w_i: X \to$

(其压缩因子分别为 $c_1, c_2, \cdots\cdots, c_n$)一起，就组成一个**迭代函数系统**，简称 IFS，记作$\{\ ; w_1, w_2, \cdots, w_n\}$，$c = \max(c_1, c_2, \cdots, c_n)$称为 IFS 的**压缩因子**.

[例 1] 考察平面欧氏空间中的三个仿射变换：

$$w_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix},$$

$$w_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 0 \\ 1/2 \end{pmatrix},$$

$$w_3\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}.$$

这三个变换是把平面上的一个正方形的每边缩小二分之一后放到三个指定的位置，如图 2-2 所示.

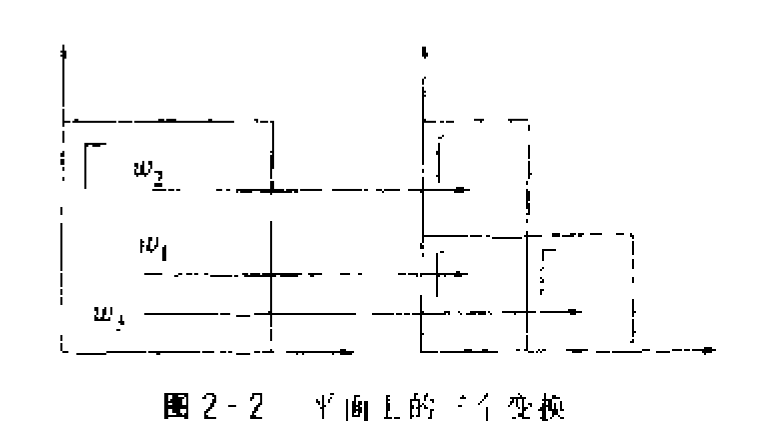


图 2-2　平面上的三个变换

显然，这三个变换都是压缩的，压缩因子均为 1/2. 因此$\{\ ; w_1, w_2, w_3\}$就是一个 IFS.

现在我们把该 IFS 对两个不同的初始图形反复进行迭代(就是把这三个变换作用到某个图形上，得到一个新图形，然后再把这三个变换作用到该新图形上，得到另一个新图形，如此反复进行)，将会得到怎样的结果呢？图 2-3 表示了对一个正方形和一个圆的

迭代过程.由于三个变换都是压缩的，每次迭代后的新图形都是缩小的，初始图形的迭代极限将成为一个点.所以，对于不论怎样的初始图形，反复迭代后将逼近于同样的极限图形.图2－3的极限图形就是图2－4所示的谢尔宾斯基三角形.它是§1例3的变形.

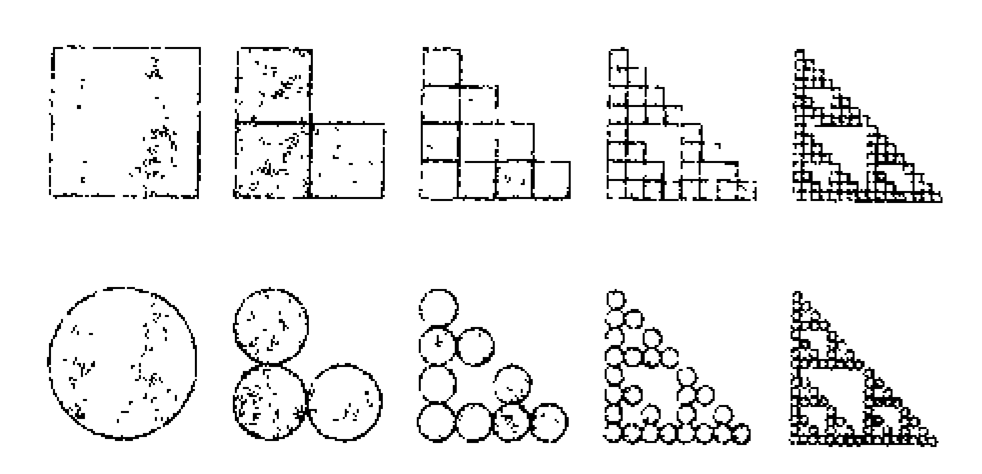


图2－3　三个变换对正方形和圆的作用

下面介绍的两个重要定理，是分形图象压缩的主要原理.

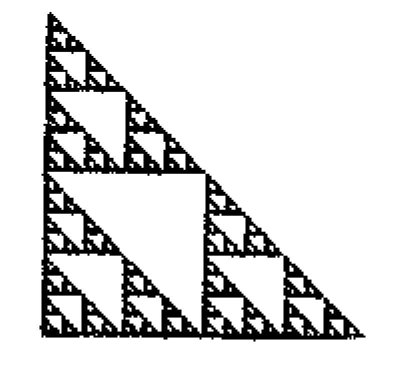


图2－4　谢尔宾斯基三角形

定理2　（压缩映射的不动点定理）设$\{X; w_1, w_2, \cdots, w_n; c\}$是$(X, d)$上的IFS，则：

(1)由下式定义的变换$W: \mathscr{H}(X) \to \mathscr{H}(X)$，即

$$W(B) = \bigcup_{i=1}^{n} w_i(B) \qquad \forall B \in \mathscr{H}(X)$$

是完备度量空间$(\mathscr{H}(X), h)$上的压缩映射，其压缩因子也是c，即

$$h(W(A), W(B)) \leqslant c \cdot h(A, B).$$

(2)压缩变换W存在唯一的不动点$\overline{A} \in \mathscr{H}(X)$，满足

$$\overline{A} = W(\overline{A}) = \bigcup_{i=1}^{n} w_i(\overline{A}); \qquad (2.8)$$

而且不动点可以通过迭代而得到，即

$$A = \lim_{n \to \infty} W^{n}(B), \qquad \forall B \in \mathcal{H}(\ ) \tag{2.9}$$

其中 $W^{1}(B) = W(B)$，而 $W^{n}(B) = W(W^{n-1}(B))$.

定理 3（拼贴定理 collage theorem） 设（ ，d）是完备的度量空间. 给定集合 $L \in \mathcal{H}(\ )$ 和数 $\varepsilon > 0$，如能选到一个 IFS{ ；w_1，w_2，…，w_n；c}（$0 \leqslant c < 1$），使

$$h\left(L, \bigcup_{i=1}^{n} w_i(L)\right) \leqslant \varepsilon,$$

则

$$h(L, \bar{A}) \leqslant \frac{h\left(L, \bigcup_{i=1}^{n} w_i(L)\right)}{1-c} \leqslant \frac{\varepsilon}{1-c}, \tag{2.10}$$

其中 h 是豪斯多夫距离，而 $\bar{A}$ 是该 IFS 的不动点（又称吸引子）.

定理 2 和定理 3 的证明可参见[4].

由上所述，一个 IFS 就是一组压缩映射，令 $W(L) = \bigcup_{i=1}^{n} w_i(L)$，由不动点定理，它确定一个唯一的吸引子. 由于唯一性，它完全由映射 W 所确定. 现在提出它的逆问题：假定我们给定某个集 L，能否找到一个 IFS，使它的吸引子恰为 L 呢？对此，到目前为止还不能解决. 但从(2.8)式和(2.10)式可以得到某些启发. (2.8)式告诉我们，不动点 $\bar{A}$ 是从复制它自身的变换 $W(\bar{A})$ 构造出来的，所以我们取给定的集 L，对它作压缩变换，然后把它们粘贴到一起以便重构 L. 不动点的唯一性是重要的，因为假定给出 L 而能找到 W（或者，假定给出 W，我们能找到集 L）使 $L = W(L)$，则一定有 $L = \bar{A}$，即 L 就是 W 的吸引子. 而方程(2.10)告诉我们，即使我们粘贴得不能使之精确地符合，在原始集 L 和粘贴后的"拼贴" $W(L)$ 之间能较好地符合，吸引子 $\bar{A}$ 将十分接近于 L.（当 c 接近 1 时，就不好了.）

[例 2] 如图 2-5(i)，取 L 是正方形. 它可由 4 个缩减 1/2 的自身复制品（压缩变换 w_1，…，w_4 如图 2-5(i)所示）精确地覆盖. 显然 $L = w_1(L) \cup w_2(L) \cup w_3(L) \cup w_4(L) = W(L)$，所以 W 的不动点 $\bar{A} = L$. 图 2-5(ii)表示取初始集为圆，以及 W 的头 3 次迭

代过程.

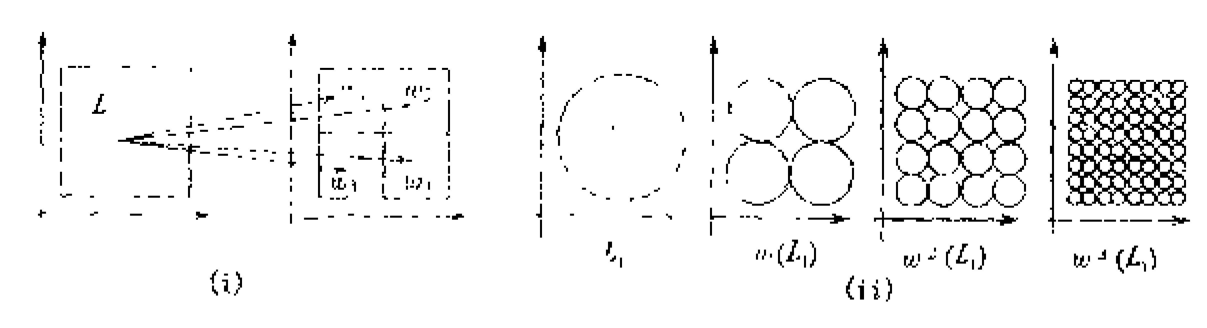
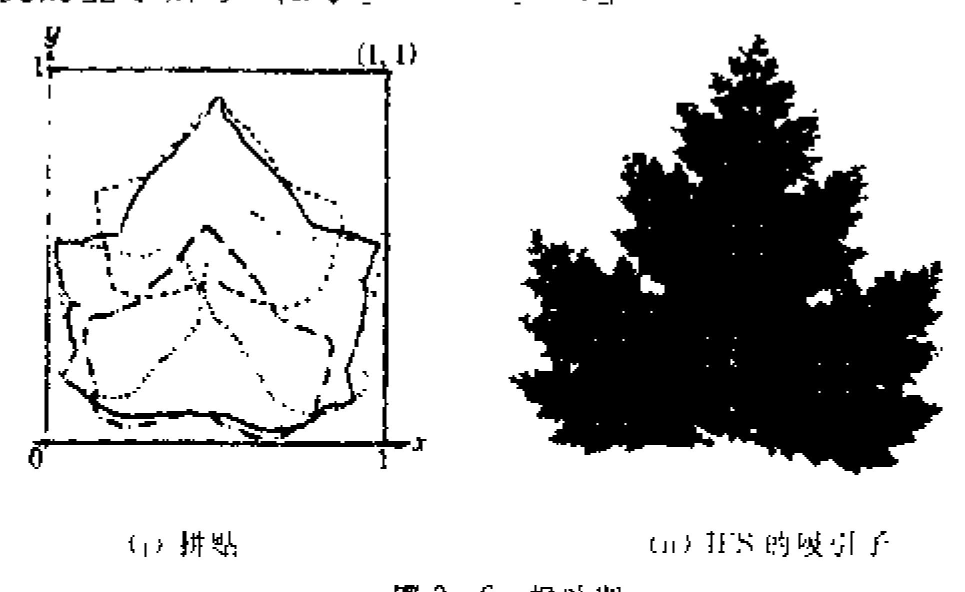


图 2-5　拼贴定理和不动点定理示例

[例 3]　枫叶的 IFS. 本例说明拼贴定理的应用.

给定的目标图形 L 可以是数字化的图象(例如黑色背景上的一片白叶),也可以是近似的多边形(例如多边形的树叶边界). 本例以枫叶作为 L,如图 2-6(i)的实线所示,设想将它显示在某一计算机的显示屏上,位于$[0,1]\times[0,1]$.



(i) 拼贴　　　　(ii) IFS 的吸引子

图 2-6　枫叶图

现在引进平面仿射变换

$$w_1\begin{bmatrix}x\\y\end{bmatrix}=\begin{bmatrix}a_1&b_1\\c_1&d_1\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}e_1\\f_1\end{bmatrix},$$

初始化的系数取为 $a_1=d_1=0.25$, $b_1=c_1=e_1=f_1=0$. 将 w_1 作用于 L,得 $\hat{w}_1(L)$,将它以不同的颜色也显示在屏幕上. 其实

$w_1(L)$ 是 L 的四分之一大小的复制品，而且它的中心很靠近(0,0)点. 现在不断地调整 w_1 的变换系数 $a_1,b_1,\cdots,f_1$，使得 $w_1(L)$ 覆盖 L 的一部分. 这样，$w_1(L)$ 就成为 L 象素的子集. 这里，重要的是 $w_1(L)$ 的大小要比 L 小，以便保证 w_1 是压缩变换. 一旦 $w_1(L)$ 的位置被确定，就引进新的仿射压缩变换 w_2，然后适当地调整 w_2 的变换系数，直至 $w_2(L)$ 覆盖住 L 象素的另一个子集(与子集 $w_1(L)$ 不同). 在 $w_1(L)$ 与 $w_2(L)$ 之间容许有重叠，但应尽可能减少重叠以便取得好的效果. 进而，再引进变换 w_3，等等. 由此即可确定一组压缩仿射变换 $\{w_1,w_2,\cdots,w_n\}$，它具有这样的性质，使给定的目标图形与集合

$$\tilde{L}=\bigcup_{i=1}^{n}w_i(L)$$

"看上去十分接近"(就是豪斯多夫距离 $h(L,\tilde{L})$ 很小)，而 n 要尽可能小. 这样确定的一组映射 $\{w_1,w_2,\cdots,w_n\}$ 就被存储起来. 而拼贴定理告诉我们，用这组映射求得的吸引子 $\tilde{A}$ 也"看上去十分接近"于目标图形 L. 如果 $L=\tilde{L}$，精确地说，如 $h(L,\tilde{L})=0$，则 $\tilde{A}=L$.

图 2-6(i)显示了对枫叶的拼贴过程，求得的 4 个仿射压缩变换的系数如表 2-1 所示，而图 2-6(ii)是这组变换的吸引子.

表 2-1 图 2-6 的 IFS 参数

	a	b	c	d	e	f
w_1	0.6	0	0	0.6	0.18	0.36
w_2	0.6	0	0	0.6	0.18	0.12
w_3	0.4	0.3	-0.3	0.4	0.27	0.36
w_4	0.4	-0.3	0.3	0.4	0.27	0.09

4. 仿射变换

在迭代函数系统(IFS)$\{X;w_1,w_2,\cdots,w_n\}$中，要求映射 w_i $(i=1,2,\cdots,n)$是压缩的. 本书都采用仿射压缩变换(如例 3).

仿射变换是一类重要的变换，n 维欧氏空间 $\mathbb{R}^n$ 中的仿射变换 $w:\mathbb{R}^n\to\mathbb{R}^n$ 具有下面的形式：

$$w(x) = A(x) + b, \qquad (x \in \mathbb{R}^n)$$

其中 A 是 $\mathbb{R}^n$ 上的线性变换(可以表示成一个 $n \times n$ 矩阵),而 b 是 $\mathbb{R}^n$ 中的一个矢量.因此,仿射变换是平移、旋转、伸缩以及反射的组合.特别地,仿射变换把球映射成椭球,把正方形映射成平行四边形,等等.显然,相似变换是仿射变换的特例.与相似变换不同的是,仿射变换在不同的方向上可有不同的伸缩比.

对于平面上的仿射变换,有

定义 4 一个变换 $w: \mathbb{R}^2 \to \mathbb{R}^2$ 如果具有下列形式:

$$w\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}, \tag{2.11}$$

其中 a、b、c、d、e 和 f 是实数,则称 w 是(2 维)**仿射变换**.这样,当它的线性部分是压缩的,该仿射变换就是压缩的,但这要依赖于所采用的度量,本书§3将会具体介绍.

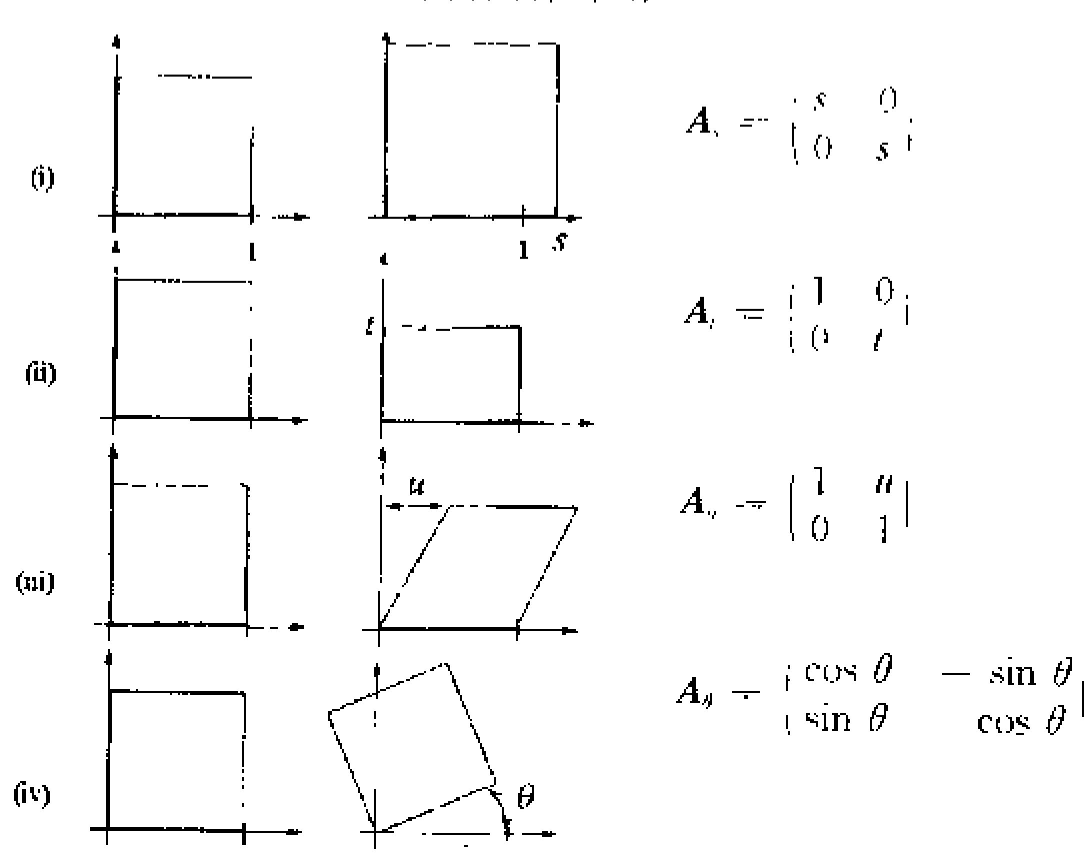


图 2.7 4 种特殊的线性变换

平面线性变换有四种特例,具有明显的几何意义,即 A_s—按比例缩放,A_t—伸长,A_u—剪切,A_θ—旋转,如图 2.7 所示.还可以

证明，任一平面仿射变换都可分解为这四种变换之乘积[1].

怎样获得我们所要求的仿射变换呢？现以下例说明.

[例4] xy 平面上的两片树叶(图2-8)，现要求出从大树叶到小树叶的仿射变换 w. 即要找到(2.11)式中的6个系数 a,b,c,d,e 和 f，使变换 w 满足

$$w(\text{大树叶})\approx\text{小树叶}.$$

首先，建立 x 和 y 坐标轴(如图示). 然后，在大树叶上选3个较有代表性的点并定出它们的坐标. 例如，取叶尖(x_1,x_2)，穗点(y_1,y_2)和连接树枝的叶柄(z_1,z_2). 再在小树叶上也取相应的3点，设其坐标分别为$(\bar{x}_1,\bar{x}_2)$，$(\bar{y}_1,\bar{y}_2)$，和$(\bar{z}_1,\bar{z}_2)$. 最后，把这些点的坐标代入(2.11)式，即可得到6个线性代数方程

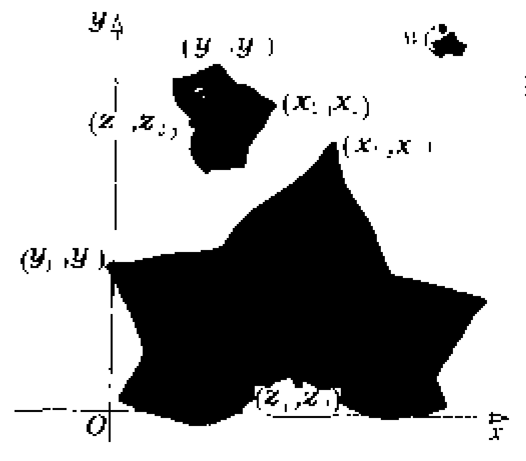


图2-8 由两片树叶确定一个仿射变换

$$\begin{cases}x_1a+x_2b+e=\bar{x}_1,\\y_1a+y_2b+e=\bar{y}_1,\\z_1a+z_2b+e=\bar{z}_1;\end{cases}$$

$$\begin{cases}x_1c+x_2d+f=\bar{x}_2,\\y_1c+y_2d+f=\bar{y}_2,\\z_1c+z_2d+f=\bar{z}_2.\end{cases}$$

由前3个方程可求得 a、b、e. 由后3个方程可求出 c、d、f. 6个系数(a,b,c,d,e,f)就确定了大小树叶之间的仿射变换.

5. 用迭代函数系统技术构造分形

当已知一个二维的IFS，即$\{R^2;w_1,w_2,\cdots,w_n\}$，其中$w_i(i=1,2,\cdots,n)$是压缩仿射变换，就可用计算机构造出相应的**分形**. 这在图象压缩技术中称为**解码**. 即从一组仿射变换(的系数)迭代出图象. 现在介绍两种算法. 一种是确定性算法，另一种是随机算法.

(1) 确定性算法：多重缩小复制机(Multiple Reduction Copying Machine，MRCM)

其实,这种方法就是普通的IFS方法.现以下例来说明.

[例5] 谢尔宾斯基垫片的IFS.

根据§1例3的谢尔宾斯基垫片的生成过程,参考本节的例1,不难求得图1-3所示谢尔宾斯基垫片IFS为{ ;w_1,w_2,w_3},其三个压缩仿射变换系数(称为IFS码)如下:

表2-2 谢尔宾斯基垫片的IFS码

	a	b	c	d	e	f
w_1	0.5	0	0	0.5	0	0
w_2	0.5	0	0	0.5	0.25	0.5
w_3	0.5	0	0	0.5	0.5	0

三个压缩仿射变换的意义如图2-9.

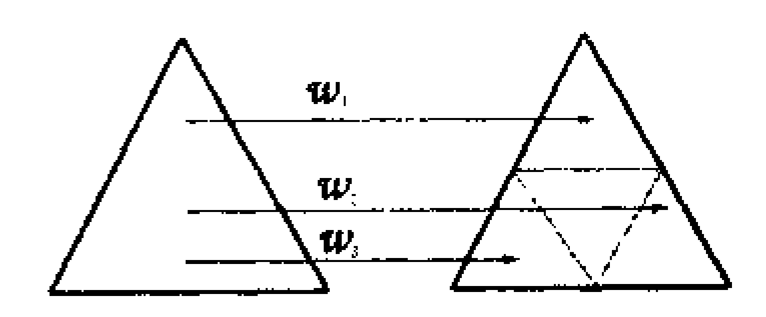


图2-9 谢尔宾斯基垫片的三个仿射变换

有了IFS后,我们任选一个紧子集A(现在是三角形,也可以是正方形或圆,等等)作为初始集,然后按

$$A_{n+1} = w_1(A_n) \cup w_2(A_n) \cup w_3(A_n) \qquad n = 0,1,2$$

逐次计算,就得出序列$\{A_i; i=0,1,2,\cdots\} \subset \mathscr{H}(\ )$.根据定理1的不动点原理,序列$\{A_i\}$收敛于(在豪斯多夫度量下)该IFS的吸引子$\overline{A}$,也就是谢尔宾斯基垫片.

对于给定的一个IFS{ ;w_1,w_2,……,w_n},算法为:

1.任选初始图形A_0,作为输入;

2.复制$w_1(A_0)$,$w_2(A_0)$,…,$w_n(A_0)$,并令$A_1 = \bigcup_{i=1}^{n} w_i(A_0)$,输出$A_1$;

3.将A_1作为输入,重复第2步,输出A_2;

……

4. 最后，$\bar{A}=\lim_{k\to\infty} A_k$，$\bar{A}$是最终图形，即为 IFS 的吸引子.

实际计算中，当第 k 步得到的图形 A_k 与下一步得到的图形 A_{k+1} 十分接近时，就把 A_k 作为最终图形.

以上的算法可以比喻为一台虚构的多重缩小图象复制机(MRCM)，如图 2-10. 这台图象复制机与普通复印机的差别，在于它具有若干个有缩小功能的透镜，每个透镜都能复制输入该机的原始图象. 透镜的缩小倍数可以按需调节，并可将缩小后的图象定位于任何规定的位置上. 因此，该机具有对输入图象作仿射变换的功能. 现按照给定的 IFS 码调节每个透镜的缩小倍数及图象的定位位置，则对任一幅原始图象，经过这种复制后输出的图象，再经过该机的反馈装置，又作为新的输入图象，反复进行处理，最后就产生出所需的分形图象. 显然，它与原始图象的形状无关.

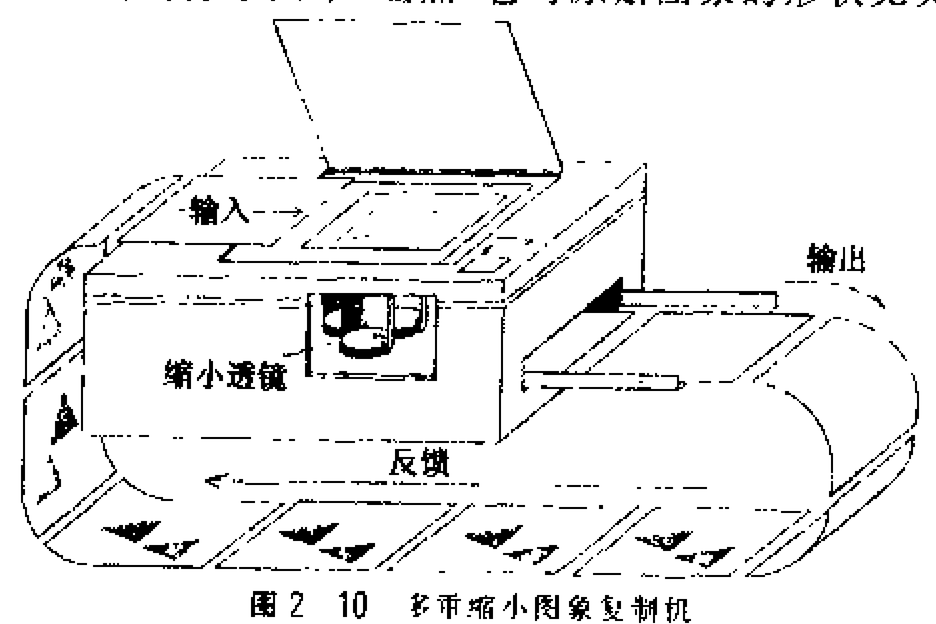


图 2-10　多重缩小图象复制机

[例 6] 蕨类植物形状的分形.

表 2-3 是某蕨类植物的 IFS 码，是由 4 个仿射变换组成的. 图 2-11 的第(i)幅矩形图是输入图象复制机的原始图象，第(ii)～(iv)幅图是复制机最初 3 次的输出图象，最后一幅就是所求的分形图象.

表2-3 蕨类植物的 IFS 码

	a	b	c	d	e	f
w_1	0	0	0	0.1[illegible]	0	0
w_2	0.8496	0.0255	-0.0255	0.8496	0	3
w_3	-0.1554	0.235	0.1958	0.1864	0	1.2
w_4	0.1554	-0.235	0.1958	0.1864	0	3

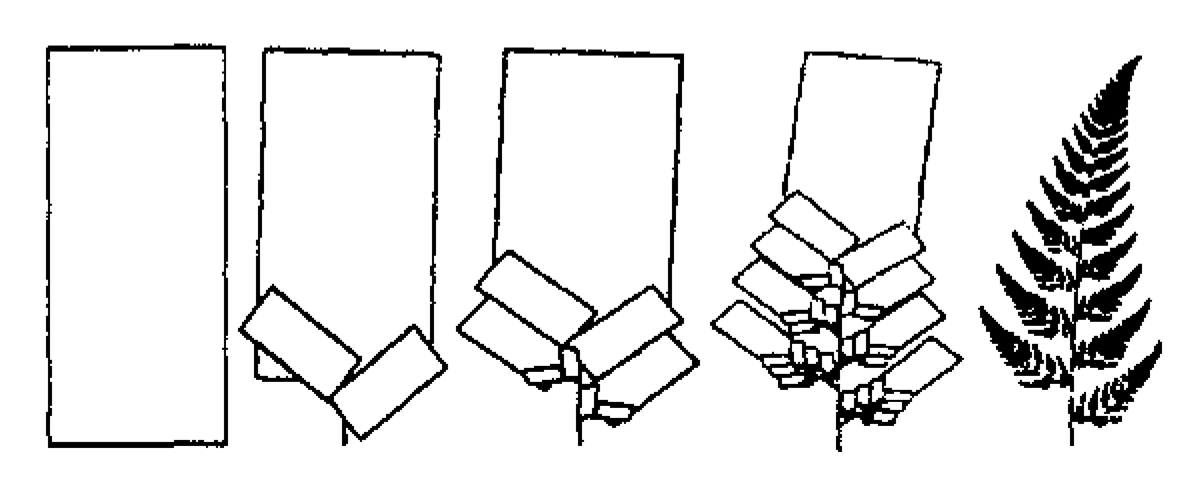


图2-11 蕨类植物分形的复制过程

[例 7] 一些经典的分形的 IFS 码.

对于§1 中讲到的一些经典的分形,根据其构造过程,即可求得其迭代函数系统的仿射变换系数(a,b,c,d,e,f)如下.

1. 康托集(§1 例 1,图 1-2,第 5 页). 容易看出,只需两个仿射变换,$w_1=(1/3,0,0,0,0,0)$,$w_2=(1/3,0,0,0,2/3,0)$,而 [illegible] = [0,1],实际上是 1 维的仿射变换.

2. 科克曲线(§1 例 2,图 1-4,第 7 页). 它的 IFS 由 4 个仿射变换组成. $w_1=(1/3,0,0,1/3,0,0)$,$w_2=(1/6,-\sqrt{3}/6,\sqrt{3}/6,1/6,1/3,0,)$,$w_3=(1/6,\sqrt{3}/6,-\sqrt{3}/6,1/6,1/2,\sqrt{3}/6)$;$w_4=(1/3,0,0,1/3,2/3,0)$.

3. 谢尔宾斯基垫片(§1 例 3,图 1-5,第 8 页). 它的 3 个仿射变换已在§2 的例 5(第 20 页)中说明. 而§2 中图 2-4 所示(第

14 页)的谢尔宾斯基三角形属于同一类，只是仿射变换 w_3 中的平移量有改变．它的 3 个仿射变换为：$w_1=(0.5,0,0,0.5,0,0)$，$w_2=(0.5,0,0,0.5,0,0.5)$，$w_3=(0.5,0,0,0.5,0.5,0)$.

(2)随机迭代算法

前已指出，当已知一个 IFS 后，总能由上述确定性算法，得到分形极限图象(即吸引子)．但是，在实际应用中，却会遇到很大困难．以本书第 21 页上例 6(蕨类植物分形)为例(见图 2-11)，初始图象是一个矩形，每完成一次迭代就要增加 4 倍的矩形，要使其极限图象显现出来大约必须完成 50 次迭代，即要计算并绘出 4^{50}($\approx 1.27\times10^{30}$)个矩形，这是目前的计算机力所不及的．实际的分形极限图象是由随机迭代算法求出的.

随机迭代算法是又一种计算 ^{2} 中迭代函数系统吸引子的方法．这种算法要在原来的 IFS 中增加一组概率，称为**带概率的迭代函数系统**(Iterated Function System with Probability)，简称 IFSP.

定义 5 一个**带概率的随机迭代函数系统**由一个 IFS{ ；w_1，w_2，…，w_n}与一组有序数{p_1，p_2，…，p_n}组成，其中

$$p_1+p_2+\cdots\cdots+p_n=1 \text{ 以及 } p_i>0, \quad (i=1,2,\cdots,n) \tag{2.12}$$

而每个概率 p_i 对应于变换 w_i.

所以，一个 IFSP 就是

$$\{ ;w_1,w_2,\cdots,w_n;p_1,p_2,\cdots,p_n\}. \tag{2.13}$$

为介绍随机迭代算法的数学原理[9]，由于在 IFSP 中含有一组概率，这就要引进新的度量空间以及相应的拼贴定理等等.

定义 6 在域 上的**测度** μ 是一个非负的实函数 μ: $\to[0,\infty)\subseteq$ ，使得任何 $A_i\in$ $(i=1,2,3,\cdots)$且 $A_i\cap A_j\neq\varnothing(i\neq j)$以及 $\bigcup_{i=1}^{\infty}A_i\in$ ，有

$$\mu\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}\mu(A_i). \tag{2.14}$$

定义 7 设(，d)是紧度量空间，将 的波莱尔(E. Borel)子

集记作 ，又设 μ 是 上的测度，则称 μ 是**波莱尔测度**. 如果 $\mu(\ )=1$，则称 μ 是**标准波莱尔测度**.

下面引入的紧度量空间 与以前的有所不同. 该空间的每个元素都是测度！

定义 8 设(,d)是紧度量空间，记 ( )是 上的标准波莱尔测度构成的集. 将 ( )上的**哈钦森度量** d_H 定义为

$$d_H(\mu,\nu)=\sup\left\{\left|\int f d\mu-\int f d\nu\right|:\right.$$

f: → 是连续的，且满足

$|f(x)-f(y)|\leqslant d(x,y)$， $\forall x,y\in$ }，

对一切 $\mu,\nu\in$ ( ). (2.15)

上述定义中，$\int fd\mu$ 和 $\int fd\nu$ 表示连续的实值函数 f 在测度上的积分，有关定义可参见[9].

定理 4 设(,d)是紧度量空间，记 ( )是 上标准波莱尔测度构成的集，又以 d_H 记其哈钦森度量，则(( ),d_H)是紧度量空间.

定理 4 的证明可见[10].

设 w: → 是连续的，可以证明[9]存在 w^{-1}: → ，于是，如果 ν 是 上的标准波莱尔测度，就有 $\nu\circ w^{-1}$: ( )→ ( ).

定义 9 设(,d)是紧度量空间，记 ( )是 上标准波莱尔测度构成的集. 又设{ ;$w_1,w_2,\cdots,w_n$;$p_1,p_2,\cdots,p_n$}是带概率的迭代函数系统，则该 IFSP 的**马尔可夫算子**(Markov operator)是函数 M: ( )→ ( )，定义为

$$M(\nu)=p_1\nu\circ w_1^{-1}+p_2\nu\circ w_2^{-1}+\cdots\cdots+p_n\nu\circ w_n^{-1},$$

$$\nu\in\ (\ ) \quad (2.16)$$

下面的定理 5 说明 M 是 ( )上的压缩映射. 为此，先介绍引理 3.

引理 3 记 M 为(2.13)式的马尔可夫算子，设 f: → 是简单函数或连续函数，令 $\nu\in$ ( )，则

$$\int fd(M(\nu)) = \sum_{i=1}^{n} p_i \int f \circ w_i d\nu. \tag{2.17}$$

证明 假设 f: → 是连续的，由定理4，$(\mathscr{P}(\ ), d_H)$是紧度量空间，因此存在一个简单函数序列$\{f_k\}$，它一致收敛于f，其中k是正整数. 容易证明

$$\int f_k d(M(\nu)) = \sum_{i=1}^{n} p_i \int_X f_k d\nu \circ w^{-1}$$

$$= \sum_{i=1}^{n} p_i \int_{w_i(X)} f_k d\nu \circ w^{-1} = \sum_{i=1}^{n} p_i \int_X f_n \circ w_i d\nu,$$

这就是序列$\left\{\int f_k d(M(\nu))\right\}$收敛于$\int fd(M(\nu))$. □

对于每个$i \in \{1,2,\cdots,n\}$和每个正整数k，$f_k \circ w_i$是简单函数，序列$\{f_k \circ w_i\}_{k=1}^{\infty}$一致收敛于$f \circ w_i$. 于是$\left\{\int f_k \circ w_i d\nu\right\}_{k=1}^{\infty}$收敛于$\int f \circ w_i d\nu$. 这样，就有$\left\{\sum_{i=1}^{n} p_i \int f_k \circ w_i d\nu\right\}_{k=1}^{\infty}$收敛于$\sum_{i=1}^{n} p_i \int f \circ w_i d\nu$.

定理5 设$(\ , d)$是紧度量空间，$\{\ ; w_1, w_2, \cdots, w_n; p_1, p_2, \cdots, p_n\}$是一个IFSP，$c \in (0,1)$是IFS的压缩因子，又设$M: \mathscr{P}(\ ) \to \mathscr{P}(\ )$是该IFSP的马尔可夫算子，则$M$是空间$(\mathscr{P}(\ ), d_H)$上的压缩映射，且压缩因子也是$c$. 即有

$$d_H(m(\nu), M(\mu)) \leqslant c d_H(\nu, \mu).$$

$$\forall \nu, \mu \in \mathscr{P}(\ ) \tag{2.18}$$

特别地，存在唯一的测度$\bar{\mu}$，满足

$$M(\bar{\mu}) = \bar{\mu}. \tag{2.19}$$

证明 记L为这样的连续函数f: → 的集合，其中的f满足$|f(x) - f(y)| \leqslant d(x,y)$，$\forall x、y \in$ ，则

$$d_H(M(\nu), M(\mu)) = \sup\left\{\int fd(M(\mu)) - \int fd(M(\nu)); f \in L\right\}$$

$$= \sup\left\{\int \sum_{i=1}^{n} p_i f \circ w_i d\mu\right.$$

$$- \int \sum_{i=1}^{n} p_i f\circ w_i d\nu ; f \in L \Big\}.$$

令 $\bar{f} = \frac{1}{c}\sum_{i=1}^{n} p_i f\circ w_i$，则 $\bar{f} \in L$. 再令

$$\bar{L} = \Big\{ \bar{f} \in L : \bar{f} = \frac{1}{c}\sum_{i=1}^{n} p_i f \circ w_i, \text{某些 } f \in L \Big\},$$

因此有

$$d_H(m(\nu), M(\mu)) = \sup\Big\{ c\int \bar{f} d\mu - c\int \bar{f} d\nu ; \bar{f} \in \bar{L} \Big\}.$$

因为 $\bar{L} \subset L$，这就推出

$$d_H(M(\nu), M(\mu)) \leqslant c\, d_H(\nu, \mu). \qquad \square$$

我们把定理 5 中的不动点 $\bar{\mu}$ 称为该 IFSP 的不变测度. 可以证明，$\bar{\mu}$ 的支撑就是 IFS$\{\ ; w_1, w_2, \cdots, w_n\}$ 的吸引子.

定理 6 （对测度的拼贴定理） 设 IFSP 是 $\{\ ; w_1, w_2, \cdots, w_n ; p_1, p_2, \cdots, p_n\}$，$\bar{\mu}$ 是它的不变测度，$c \in (0, 1)$ 是 IFS 的压缩因子；又设 $M: \mathscr{P}(\ ) \to \mathscr{P}(\ )$ 是该 IFSP 的马尔可夫算子，$\nu \in \mathscr{P}(\ )$，则

$$d_H(\nu, \bar{\mu}) \leqslant \frac{d_H(\nu, M(\nu))}{1 - c}. \tag{2.20}$$

证明 根据哈钦森度量的三角不等式、不变测度的定义(2.19)式以及定理 5 中的(2.18)式，可得

$$\begin{aligned} d_H(\nu, \bar{\mu}) &\leqslant d_H(\nu, M(\nu)) + d_H(M(\nu), \bar{\mu}) \\ &= d_H(\nu, M(\nu)) + d_H(M(\nu), M(\bar{\mu})) \\ &\leqslant d_H(\nu, M(\nu)) + c\, d_H(\nu, \bar{\mu}). \end{aligned} \qquad \square$$

下面的定理和它的推论声称某些事件的出现"具有概率 1". 这就是埃尔顿(J. Elton)的遍历定理.

定理 7 设$(\ , d)$是紧度量空间，一个 IFSP 是$\{\ ; w_1, w_2, \cdots, w_n ; p_1, p_2, \cdots, p_n\}$. 记 $\{x_k\}_{k=0}^{\infty}$ 是由随机迭代算法产生的 IFS 的一条轨道(起始点为 x_0)，即

$$x_k = w_{\sigma_k} \circ w_{\sigma_{k-1}} \circ \cdots \circ w_{\sigma_1}(x_0), \tag{2.21}$$

其中映射 $w_{\sigma_1}, w_{\sigma_2}, \cdots, w_{\sigma_k}$ 是按照概率 $p_1, p_2, \cdots, p_n$ 独立选取的，$k=1,2,3,\cdots$. 现设 $\overline{\mu}$ 是 IFSP 的唯一不变测度，则下式具有概率 1（即对所有码序列 $\sigma_1, \sigma_2, \cdots$，除了概率为 0 的序列集）：

$$\lim_{k\to\infty}\frac{1}{k+1}\sum_{i=0}^{k} f(x_i) = \int_{\mathscr{X}} f(x)d\mu(x) \qquad (2.22)$$

对一切连续函数 $f: \mathscr{X} \to \mathscr{R}$ 和一切 x_0.

定理 7 的证明可见[11].

定理 7 的推论 设 B 是 $\mathscr{X}$ 的波莱尔子集，且 μ(B 集的边界)$=0$，又令

$$\mathscr{N}(B,k) = \text{在交集 } \{x_0, x_1, x_2\cdots, x_k\} \cap B \text{ 中的点数},$$

对 $k=0,1,2,\cdots$，则下式具有概率 1：

$$\mu(B) = \lim_{k\to\infty}\left\{\frac{\mathscr{N}(B,k)}{k+1}\right\} \text{ 对一切从 } x_0 \text{ 开始的点}. \qquad (2.23)$$

这样，B 的测度与迭代的步数成比例. 当执行随机迭代算法时，它就产生 B 中的点.

有了以上的理论基础，就保证了随机迭代算法的成功.

随机迭代算法 设 IFSP 是 $\{\mathscr{X}; w_1, w_2, \cdots, w_n; p_1, p_2, \cdots, p_n\}$，其中 $p_i>0$ 并已分配给 w_i，对 $i=1,2,\cdots,n$；而且 $\sum_{i=1}^{n} p_i = 1$. 现任取 $x_0 \in \mathscr{X}$，然后反复地、独立地选取

$$x_k \in \{w_1(x_{k-1}), w_2(x_{k-1}), \cdots, w_n(x_{k-1})\},$$

$$\text{对 } k = 1,2,3,\cdots \qquad (2.24)$$

其中事件 $x_k = w_i(x_{k-1})$ 的概率是 p_i. 这样就构造出点列 $\{x_k: k=0, 1, 2, \cdots\} \subset \mathscr{X}$.

该点列 $\{x_k\}_{k=0}^{\infty}$（以 100% 的概率）"收敛到"IFS 的吸引子. 而且，如果点 x_0 刚好就是不动点之一，则该序列中的所有点都位于吸引子上.

算法的具体步骤如下（计算机作图）：

1. 不妨取点 $x=(0,0)$.

2. 在数列$\{1,2,\cdots,n\}$中，以概率 p_k 选出数 k.

可以想象成有一颗 n 个面的骰子，每个面上分别标上数字 1，2，…，n，且第 i 个面出现的概率为 p_i. 现在掷这样的骰子，设出现第 k 个面，就选到了数字 k.

3. 把第 k 个压缩仿射变换 w_k 作用到点 x 上，得到新的点，记作 $\mathrm{new}x=w_k(x)$.

4. 把点 $\mathrm{new}x$ 仍记为 x. 即令 $x:=\mathrm{new}x$.

5. 画点 x.

6. 回到第 2 步.

根据以上步骤，即可编写有关的计算程序. 在实际作图时，刚开始的 300 个点可以不画. 通常，1～2 秒内就可画出成千上万个点，一幅分形图就会显现出来.

关于概率 $p_i(i=1,2,\cdots,n)$，从理论上来说，不论怎样给定它们的值，都不会影响最终结果；但是，在实际执行时，这会影响到计算速度. 通常，可由下列公式近似地确定这些概率值：

$$p_i\approx\frac{|a_id_i-b_ic_i|}{\sum\limits_{i=1}^{n}|a_id_i-b_ic_i|}\qquad i=1,2,\cdots,N \quad (2.25)$$

也就是使 p_i 正比于第 i 个压缩仿射变换 w_i 中齐次线性变换系数构成的四边形面积.

文献[12]及其相应的软件 Desktop Fractal Design System 就是按照随机迭代算法计算出多种多样的分形.

［例 8］ 蕨类植物的 IFSP 码及其随机迭代算法的计算结果.



图 2-12 随机迭代算法的结果

表 2-4 蕨类植物的 IFSP 码

	a	b	c	d	e	f	p
w_1	0	0	0	0.16	0	0	0.01
w_2	0.85	0.04	-0.04	0.85	0	1.6	0.85
w_3	0.2	-0.26	0.23	0.22	0	1.6	0.07
w_4	-0.15	0.28	0.26	0.24	0	0.44	0.07

图 2-12 表明在执行随机迭代算法过程中，当迭代次数增加时出现的 5 幅图. 随机落下的点逐渐显现出 IFS 的吸引子(可与本书第 23 页的图 2-11 相比较).

§ 3 分形图象压缩的基本原理

现在介绍如何把分形理论(主要是 IFS 理论)应用于图象压缩.

图象压缩的分形方法是 1988 年由巴恩斯利首先提出并实现的. 近十年来，已获得极为迅速的发展.

1. 图象模型

首先要说明，前面所讨论的"图象"，都是指由平面上的子集所表示的黑-白图象，黑色就代表子集中的点. 而现在要讨论带灰度的图象，就像单色照片那样的图. 为了讨论图象压缩，就要建立数学模型. 常用的灰度图象模型有 3 种：测度空间、象素数据和函数，现分别介绍如下.

i. 当以测度作为图象的模型时，就是把图象表示为平面上的一种测度 μ. 这时，明暗度就能由平面子集 A 上的度量，即

$$\mu(A) = \int_A d\mu$$

来表示. 在 § 2 中介绍的随机迭代算法，就是采用这种模型. 更多的应用可见[13].

ii. 在象素数据模型中，把图象表示为离散象素的集合. 每个象素具有的离散值(通常范围为 0 至 255)代表一个灰阶. 每个象素的比特数(用来存储象素值)确定了灰度的分辨率，而总的象素数目就确定了图象的分辨率. 计算机只能在这种表示形式下工作. 本书的第 2 章和第 3 章都基于这种表示.

一幅包含有 $M\times N$ 个象素的图象，可以想象为 $n=MN$ 维空间中的一个向量. 当略去象素值的离散特性时，这空间就是 $\mathbb{R}^n$. 我们把 $\mathbb{R}^n$ 上的普通范数(即 p 范数)定义为

$$\|\boldsymbol{x}\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}. \tag{3.1}$$

由此，可定义两个向量间的度量为

$$d_p(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x}-\boldsymbol{y}\|_p, \qquad \forall\, \boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n \tag{3.2}$$

特别地，当 $p=2$ 时，2 范数又称为 l^2 范数，由它导出的度量就称为**均方根度量**，简称 rms. 于是，设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 和 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$ 是两幅图象，则它们的均方根度量由下式给出：

$$d_{rms}(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x}-\boldsymbol{y}\|_2 = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}. \tag{3.3}$$

iii. 现在讨论函数模型，这是本节采用的模型.

可把函数 $f:\mathbb{R}^2\to\mathbb{R}$ 看作是具有无限分辨率的一幅映象(图象). 因为实际映象总是在有限范围内的，我们取 f 的定义域为单位正方形 $I^2=\{(x,y):0\leqslant x,y\leqslant 1\}$，而 f 的值域为 $I=\{x:0\leqslant x\leqslant 1\}$. f 的值代表灰级.

所谓在函数空间中的 L^p 度量，是

$$d_p(f,g) = \|f-g\|_p = \left(\int_{I^2}|f-g|^p d\boldsymbol{x}\right)^{1/p}. \tag{3.4}$$

于是，均方根度量就相当于 L^2 度量.

还有一种使用上很方便的**上界度量**，即

$$d_{sup}(f,g) = \sup_{(x,y)\in I^2}|f(x,y)-g(x,y)|. \tag{3.5}$$

上式表示，找到一个点 $(x,y)\in I^2$，在该点上 f 和 g 的差别最大，就

把这个差值定义为 f 和 g 之间的距离. 这种度量很容易计算, 但不大符合实际. 例如, 有两个函数除一点以外是恒等的, 则由(3.5)式, 这两个函数间的距离可以很大. 其实这是不合理的.

函数空间上的度量的定义还有多种, 但在图象压缩理论中真正关心的是迭代过程的收敛性. 对于函数说来, 点点收敛和几乎处处收敛是两个重要概念. 所谓 f_n **点点收敛于** f, 意即对每个确定的 x 都有 $f_n(x)\to f(x)$. 而 f_n **几乎处处收敛于** f, 意即除了一个零测度集以外, f_n 点点收敛于 f. 由此可见, 关于函数序列收敛性的有关概念, 与函数空间的度量无关, 而只与函数值域上的度量有关. 因此, 可能求得了一个收敛的图象序列 f_n 而完全不讨论某种度量.

最后, 给定一个函数 $f: \mathbb{R}^2\to\mathbb{R}$, f 的**图**就是指由全体三元点 $(x, y, f(x, y))$ 所组成的 $\mathbb{R}^3$ 中的一个子集. 严格说, f 的图与 f 是不同的, 但两者包含有同样的信息, 有时就不加区别了.

2. $\mathbb{R}^2$ 上的压缩仿射变换

对 $\mathbb{R}^2$ 上的仿射变换(2.11), 现将其齐次线性变换的系数矩阵记作 $\boldsymbol{A}$, 即

$$\boldsymbol{A}=\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

设我们选取 $\mathbb{R}^2$ 上的范数 $\|\cdot\|$, 则对于 $\boldsymbol{x}\in\mathbb{R}^2$, 定义 $\boldsymbol{A}$ 的范数为

$$\|\boldsymbol{A}\| = \sup_{\boldsymbol{x}\neq 0}\frac{\|\boldsymbol{A}\boldsymbol{x}\|}{\|\boldsymbol{x}\|}. \tag{3.6}$$

当 $\|\boldsymbol{A}\|<1$ 时, 就称仿射变换(2.11)是**压缩**的.

对于 p 范数, 如取 $p=2$, 可以证明[14]

$$\|\boldsymbol{A}\|_2 = \text{矩阵 } \boldsymbol{A}^T\boldsymbol{A} \text{ 的最大特征值的平方根}.$$

具体地, 可算得 $\|\boldsymbol{A}\|_2$ 为

$$\left\{\frac{1}{2}\left[a^2+b^2+c^2+d^2+\left\{(a^2+b^2+c^2+d^2)-4(ad-bc)^2\right\}^{\frac{1}{2}}\right]\right\}^{\frac{1}{2}}. \tag{3.7}$$

例如，§2 中例 5 的谢尔宾斯基垫片（第 19 页），其三个仿射变换的系数见表 2－2，均有 $a=d=0.5$，$b=c=0$，代入（3.7）式，可算得 $\|A\|=\frac{1}{2}<1$，可见这三个变换都是压缩的.

3. 带映射的局部迭代函数系统（LIFSM）

现在讨论怎样对现实生活中的图象，例如图 3－1 所示少女列娜（Lenna）像的原始图象（256×256 象素），怎样用分形原理进行压缩？这类图象与§2 中用 IFS 方法处理的分形图有两点明显不同. 第一，它们往往是彩色的，至少有灰度. 第二，它们不包含自相似性. 例如对谢尔宾斯基垫片（§2 例 5），能找到 3 个压缩仿射变换把自身（整体）变换到局部，而列娜的脸就不能由自身的压缩仿射变换组合而成. 所幸这些困难能够解决. 首先，本节开头已指出，只要采用图象的函数模型，即增加考虑一个映射 $f: X \to R$，函数 $f(x,y)$ 就可代表每个位置 (x,y) 上的灰级，而由此也就容易处理彩色图象了（这将在第 3 章中介绍）. 其次，虽然没有整体与局部的自相似性，但经验显示，现实生活中的图象却存在局部之间的自相似性. 图 3－2 显示了从列娜像上取出的一些部位，那里具有不同比例的相似性：她肩上一部分与其重叠的一个小部位几乎是相同的，镜子里帽子的象的一部分与她帽子的一部分是相似的（经过旋转和改变亮度以后）. 可见局部自相似性与图 2－11 表示的自相似不同之处在于现在的图象是由本身的许多"部分"在适当的变换下复制构成，这些"部分"并不是它们自身在仿射变换下的恒等复制品，而是有误差的. 这样就可构成局部迭代函数系统的方法，它完全是 IFS 方法的推广，只是对每个变换 w_i 的定义域仅为 X 的一部分 D_i.

定义 1 设 (X,d) 是完备度量空间，而 $D_i \subset X$（其中 $i=1,2,\cdots,n$）. 局部迭代函数系统（简称 LIFS）就是下列压缩映射集：

$$\{w_i: D_i \to X,\quad (i=1,2,\cdots,n)\}. \tag{3.8}$$

现在考察一幅平面灰度图象，把 $f(x,y)$ 称为每个点 (x,y) 上的灰





图 3-1 原始的列娜像

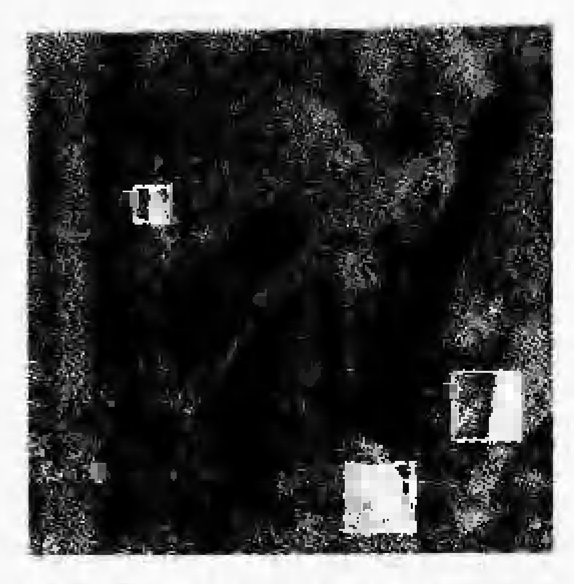


图 3-2 某些自相似的"部分"

级。为了简单起见，设 $(x,y)\in I^2=\{(u,v):0\leqslant u,v\leqslant 1\}$，及 $f(x,y)\in I=[0,1]$。图象的空间 $F=\{f:I^2\rightarrow I\}$ 容许 f 的值域为 I。这样，图象的和与差仍有定义。若采用(3.5)式定义的上界度量，则有下面的定理。

定理 1 空间 $(F, d_{\sup})$ 是完备度量空间。

定理 1 的证明见[14]，这里从略。

为了考虑灰度，我们对 2 维的平面映射上再加从级的映射作为第 3 维，就构成**带映射的局部迭代函数系统**(Local Iterated Function System with grey level Maps，简称 LIFSM)。

我们把 I^2 划分为许多块。设 $D_1,D_2,\cdots,D_n$ 和 $R_1,R_2,\cdots,R_n$ 是 I^2 的子集，分别称为**定义域块**(或称**父块**)和**值域块**(又称**子块**)。又设 $v_1,v_2,\cdots,v_n:I^3\rightarrow I^3$ 是一组映射，如果对每个 v_i 的定义域加以限制，即有

定义 2 设

$$w_i=v_i|_{D_i\times I}:\quad D_i\times I\rightarrow R_i\times I,\quad (i=1,2,\cdots,n)\tag{3.9}$$

即把 w_i 限制定义在父块 D_i 上，则称 $\{w_1,w_2,\cdots,w_n\}$ 为 IFSM。

"把 w_i 作用在 f 上"定义为 $w_i(f)=w_i(x,y,f(x,y))$，并把它看作 I^2 上一个函数的图。而要做到这点，就需要引进"铺覆"的概

念.

定义 3 上述映射 $w_1, w_2, \cdots, w_n$，如有对一切 $f \in \mathscr{F}$，使 $\bigcup_{i=1}^{n} w_i(f) \in \mathscr{F}$，则称"$w_1, w_2, \cdots, w_n$ **铺覆** I^2".

定义 3 告诉我们，当把 w_i 作用到父块 D_i 上的图象 $f \cap (D_i \times I)$ 上时，必须得到覆盖子块 R_i 上的函数的图，而且 $I^2 = \bigcup_{i=1}^{n} R_i$，见图 3-3. 所以，铺覆条件就意味着并 $\bigcup_{i=1}^{n} w_i(f)$ 产生覆盖于 I^2 上的函数的图，而且各个 R_i 互不重叠.

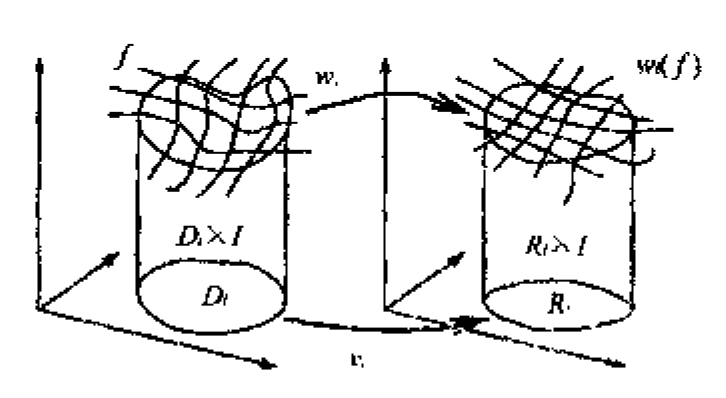


图 3-3 w_i 把 D_i 上的图映射到 R_i 上

就像 §2 的 IFS 方法一样，现在要建立空间$(\mathscr{F}, d_{\sup})$上的压缩算子；应用不动点定理，求出吸引子.

定义 4 设映射 $w: \mathbb{R}^3 \to \mathbb{R}^3$，记$(x', y', z_1') = w(x, y, z_1)$和$(x', y', z_2') = w(x, y, z_2)$. 如果存在一个正实数 $s < 1$，使

$$|z_1' - z_2'| \leqslant s|z_1 - z_2|, \tag{3.10}$$

而且 x' 和 y' 关于 z_1 或 z_2 是独立的，对一切 x, y, z_1, z_2 成立，则称映射 w 是 **z-压缩的**，并称 s 是 w 的 **z-压缩因子**.

定理 2 如果 $w_1, w_2, \cdots, w_n$ 是 z-压缩的，则

$$W = \bigcup_{i=1}^{n} w_i \tag{3.11}$$

是空间$(\mathscr{F}, d_{\sup})$的压缩映射.

证明 设 w_i 的 z-压缩因子为 s_i，令 $s = \max(s_1, s_2, \cdots, s_n)$.

$d_{\sup}(W(f), W(g)) = \sup\{|W(f)(x, y)$

$$-W^{r}(g)(x,y)|:(x,y)\in I^2\}$$

$$=\sup\{|w_i(x,y,f(x,y))-w_i(x,y,g(x,y))|$$

$$\text{的 } z \text{ 分量}:(x,y)\in D_i,\ i=1,2,\cdots,n\}$$

$$\leqslant\sup\{s_i|f(x,y)-g(x,y)|:i=1,2,\cdots,n\}$$

$$\leqslant\sup\{s|f(x,y)-g(x,y)|\}$$

$$=s\sup\{|f(x,y)-g(x,y)|\}$$

$$\leqslant sd_{\sup}(f,g).$$

上面证明中用到 w_i 的坐标 z 与 x、y 是独立的. □

虽然 LIFSM 的定义没有对其中的变换有规定，但为了方便，我们仍将采用仿射变换. 现设 z 方向压缩的仿射变换 w_i 有如下形式：

$$w_i\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}a_i & b_i & 0\\c_i & d_i & 0\\0 & 0 & s_i\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}+\begin{bmatrix}e_i\\f_i\\o_i\end{bmatrix},\tag{3.12}$$

其中 $s_i<1$，则由定理 2，$W=\bigcup\limits_{i=1}^{n}w_i$ 是压缩算子（在 $d_{\sup}$ 度量下）.

(3.12)式表示的变换 w_i，实际上是由两部分组成的. 为此，可记

$$v_i(x,y)=\begin{bmatrix}a_i & b_i\\c_i & d_i\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}e_i\\f_i\end{bmatrix}.\tag{3.13}$$

对于图象 f，因为 $w_i(f)=w_i(x,y,f(x,y))$，所以平面仿射变换 $v_i(x,y)$ 确定了把分割的定义域块映射到它的复制品上，而 s_i 和 o_i 则确定了变换的对比度和亮度. 因此，每个变换 w_i 都限制于定义域 $D_i\times I$ 上. 这样，w_i 仅仅应用于定义域块 D_i 之上的那部分图象. 即有

$$v_i(D_i)=R_i.\tag{3.14}$$

而铺覆条件保证当计算 W 时再次得到一个函数，才能使迭代继续下去. 由 W 在 F 上的压缩性、根据压缩映射的不动点定理（§2 定理 2），在 F 上存在唯一的不动点 $\bar{f}$，并可通过对任意初始函数

f_0 的迭代而得到：$f_0, W(f_0), W(W(f_0)), \cdots, \bar{f}$；而 $W(\bar{f})=\bar{f}$.

4. 利用 LIFSM 对灰度图象的编码

假设给定灰度图象 f，要求对它进行编码，以达到压缩的目的. 也就是要求 LIFSM 的逆问题：给定 f，如何能找到映射 W，使它的吸引子 $\bar{f}=f$. 这是尚未解决的问题. 但我们可降低要求，可以找到图象 $f'\in F$，使 $d(f,f')$ 达到最小，而 $\bar{f'}=f$. 因为 $\bar{f}=W(\bar{f})=\bigcup_{i=1}^{n} w_i(\bar{f})$，所以只要寻找父块 $D_1, D_2, \cdots, D_n$ 和相应的变换 $w_1, w_2, \cdots, w_n$（它们是 z－压缩的，而且铺覆 I^2），使

$$f\approx W(f)=\bigcup_{i=1}^{n} w_i(f). \tag{3.15}$$

而这可在拼贴定理（§2 定理 3）的帮助下完成.

编码过程如下：将 f 的定义域 I^2 划分为子块集 $R_1, R_2, \cdots, R_n$. 对每一个子块 R_i 寻找父块 $D_i\subset I^2$ 和变换 $w_i: D_i\times I\to I^3$，使 $w_i(f)$ 尽量接近 $f\cap(R_i\times I)$. 即距离

$$d(f\cap(R_i\times I), w_i(f))\qquad i=1,2,\cdots,n \tag{3.16}$$

达到最小，以上过程如图 3－4 所示. 由（3.15）式就得到 W，选得的 $w_1, w_2, \cdots, w_n$ 必须保证 W 是压缩的，这样就完成了编码. 由于只要储存形如（3.12）的 w_i 的系数，以及 D_i 的位置和从 D_i 到 R_i 的变换信息，这些信息量大大少于原始图象的信息量，这就实现了压缩的目的.

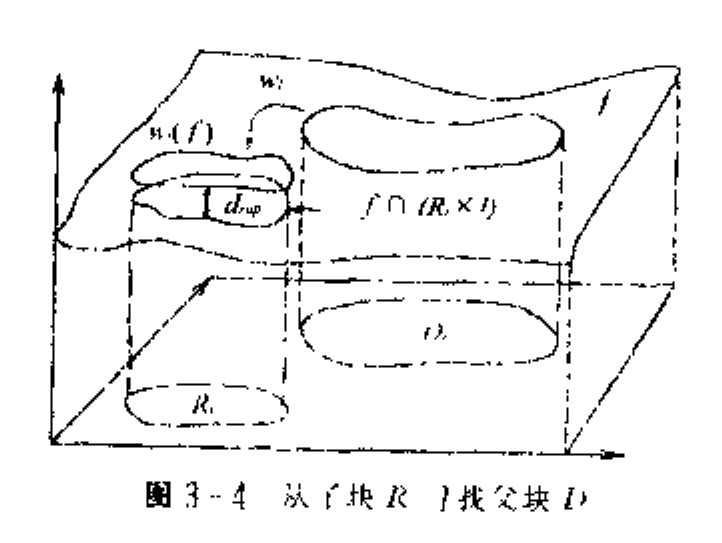


图 3－4 从子块 R_i 寻找父块 D_i

［例 1］ 一幅大小为 256×256 象素的图象，每个象素的灰级为［0，255］中一个数（表示从黑到白）. 设 $R_1, R_2, \cdots, R_{1024}$ 是 8×8 象素的非重叠子块，而父块集 $\{D_n\}$ 是所有 16×16 象素小方块（可重叠）的集合，即共包含 241

$\times 241=58081$ 个小方块. 对每一个 R_i,根据极小化方程(3.16)在父块集$\{D_n\}$中寻找一个父块 D_i,也就是找到了图象的一个"部分"使之与 R_i 上的图象十分相像,并称该父块 D_i 覆盖了子块 R_i. 有 8 种方法(见第 2 章§6)把一个方块映射到另一个,所以对 1024 个中的每个子块要比较 $8\times 58081=464648$ 个方块. 还有,父块集中每个方块的象素数目是每个 R_i 的 4 倍,所以当我们对方程(3.16)作极小化时可以仅从 D_i 中选一个 2×2 子块进行,或者对 D_i 的每个 2×2 子块平均后对应于 R_i 的 1 个象素.

对方程(3.16)式极小化的意义. 如上所述要选到好的父块 D_i,并由(3.14)式求得 v_i 的 6 个系数外,还要求对(3.12)式 w_i 选到好的 s_i 和 o_i,即好的对比度和亮度设置. 就是对父块集中的每块 D_K 用最小二乘法计算 s_K 和 o_K,然后选最小均方根大的作为 D_i.

选定 D_i 后,与相应的 s_i 和 o_i 一起,就确定了形如(3.12)式的仿射变换 w_i. 一旦全部变换 $w_1,w_2,\cdots,w_{1024}$ 找到,编码即告完成.

用 LIFSM 方法,图象的解码很简单,就是由迭代方法求出吸引子 $\bar{f}$. 先令 $W=\bigcup_{i=1}^{1024} w_i$,再取初始图象 f_0 为全白的 I^2,然后求 $W(f_0)$、$W(W(f_0))$,…,直至连续两幅图象已十分接近,就求得了 $\bar{f}\approx f$.

对图 3-1 的列娜像,计算结果为:原始图象占 65536 比特,而全部变换只占 3968 比特,压缩比为 16.5:1.

关于对本方法的改进和实验结果,将在第 3 章中介绍.

[例 2] 设 $I=[0,1]\subset$ 直线 L. 一组压缩仿射变换为:$v_1(x)=\frac{1}{3}x$,$v_2(x)=\frac{x}{3}+\frac{1}{3}$,$v_3(x)=\frac{x}{3}+\frac{2}{3}$. 设灰度映射是 $\phi_1(z)=\frac{1}{2}z$,$\phi_2(z)=\frac{1}{2}$,$\phi_3(z)=\frac{1}{2}z+\frac{1}{2}$. 试用带映射的迭代函数系统方法,求其吸引子 $\bar{f}(x)$.

例 1 是 LIFSM 的逆问题,即对给定图象进行编码,求出一组仿射变换$\{w_1,w_2,\cdots,w_{1024}\}$. 例子着重说明了寻找父块与子块相

匹配的方法．而例 2 是正问题，即进行解码，着重说明怎样通过迭代而获得吸引子．

现在，包含灰度映射的 3 个仿射变换（相应于(3.12)式）为

$$w_1\begin{pmatrix}x\\z\end{pmatrix}=\begin{pmatrix}\frac{1}{3}&0\\0&\frac{1}{2}\end{pmatrix}\begin{pmatrix}x\\z\end{pmatrix},$$

$$w_2\begin{pmatrix}x\\z\end{pmatrix}=\begin{pmatrix}\frac{1}{3}&0\\0&0\end{pmatrix}\begin{pmatrix}x\\z\end{pmatrix}+\begin{pmatrix}\frac{1}{3}\\\frac{1}{2}\end{pmatrix},$$

$$w_3\begin{pmatrix}x\\z\end{pmatrix}=\begin{pmatrix}\frac{1}{3}&0\\0&\frac{1}{2}\end{pmatrix}\begin{pmatrix}x\\z\end{pmatrix}+\begin{pmatrix}\frac{2}{3}\\\frac{1}{2}\end{pmatrix},$$

则空间 $\mathcal{F}=\{f:[0,1]\to\mathbb{R}\}$ 上的算子

$$W=w_1\cup w_2\cup w_3,$$

且有

$$\begin{aligned}W(f)&=W(x,f(x))\\&=w_1(x,f(x))\cup w_2(x,f(x))\cup w_3(x,f(x)).\end{aligned}$$

现取初始函数 $f_0(x)=0$，代入上式，得

$$\begin{aligned}W(f_0)&=w_1(x,0)\cup w_2(x,0)\cup w_3(x,0)\\&=\begin{pmatrix}\frac{x}{3}\\0\end{pmatrix}\cup\begin{pmatrix}\frac{1}{3}+\frac{x}{3}\\\frac{1}{2}\end{pmatrix}\cup\begin{pmatrix}\frac{2}{3}+\frac{x}{3}\\\frac{1}{2}\end{pmatrix},\qquad x\in[0,1]\end{aligned}$$

即得 $f_1(x)$ 为

$$f_1(x)=\begin{cases}0, & \text{当 } x\in\left[0,\frac{1}{3}\right) \text{ 时；}\\ \frac{1}{2}, & \text{当 } x\in\left[\frac{1}{3},1\right] \text{ 时．}\end{cases}$$

再作第 2 次迭代：

$$
\begin{aligned}
W(f_1) &= w_1(x,f_1) \cup w_2(x,f_1) \cup w_3(x,f_1) \\
&= \begin{pmatrix} \frac{x}{9} \\ 0 \end{pmatrix} \cup \begin{pmatrix} \frac{1}{9}+\frac{x}{9} \\ \frac{1}{4} \end{pmatrix} \cup \begin{pmatrix} \frac{2}{9}+\frac{x}{9} \\ \frac{1}{4} \end{pmatrix} \\
&\quad \cup \begin{pmatrix} \frac{1}{3}+\frac{x}{9} \\ \frac{1}{2} \end{pmatrix} \cup \begin{pmatrix} \frac{4}{9}+\frac{x}{9} \\ \frac{1}{2} \end{pmatrix} \cup \begin{pmatrix} \frac{5}{9}+\frac{x}{9} \\ \frac{1}{2} \end{pmatrix} \\
&\quad \cup \begin{pmatrix} \frac{2}{3}+\frac{x}{9} \\ \frac{1}{2} \end{pmatrix} \cup \begin{pmatrix} \frac{7}{9}+\frac{x}{9} \\ \frac{3}{4} \end{pmatrix} \cup \begin{pmatrix} \frac{8}{9}+\frac{x}{9} \\ \frac{3}{4} \end{pmatrix}, \\
&\qquad\qquad\qquad\qquad x \in [0,1]
\end{aligned}
$$

即 $f_2(x)$ 为

$$
f_2(x) = \begin{cases} 0, & 当\ x \in \left[0, \frac{1}{9}\right) 时; \\ \frac{1}{4}, & 当\ x \in \left[\frac{1}{9}, \frac{1}{3}\right) 时; \\ \frac{1}{2}, & 当\ x \in \left[\frac{1}{3}, \frac{7}{9}\right) 时; \\ \frac{3}{4}, & 当\ x \in \left[\frac{7}{9}, 1\right] 时. \end{cases}
$$

如此继续迭代，每个 $f_k(x)$ 都是阶梯函数($k=1,2,3,\cdots$)，它的极限(也就是 W 的吸引子)$\bar{f}(x)$，称为"魔鬼楼梯函数"；对 $X=[0,1]$ 中除了 $\mathscr{K}$ 集外每点都存在导数，而 $\mathscr{K}$ 是[0,1]上的三元康托集(见 §1 例 1).

函数 $f_0(x)$、$f_1(x)$、$f_2(x)$ 和 $\bar{f}(x)$ 的图形，见图 3-5.

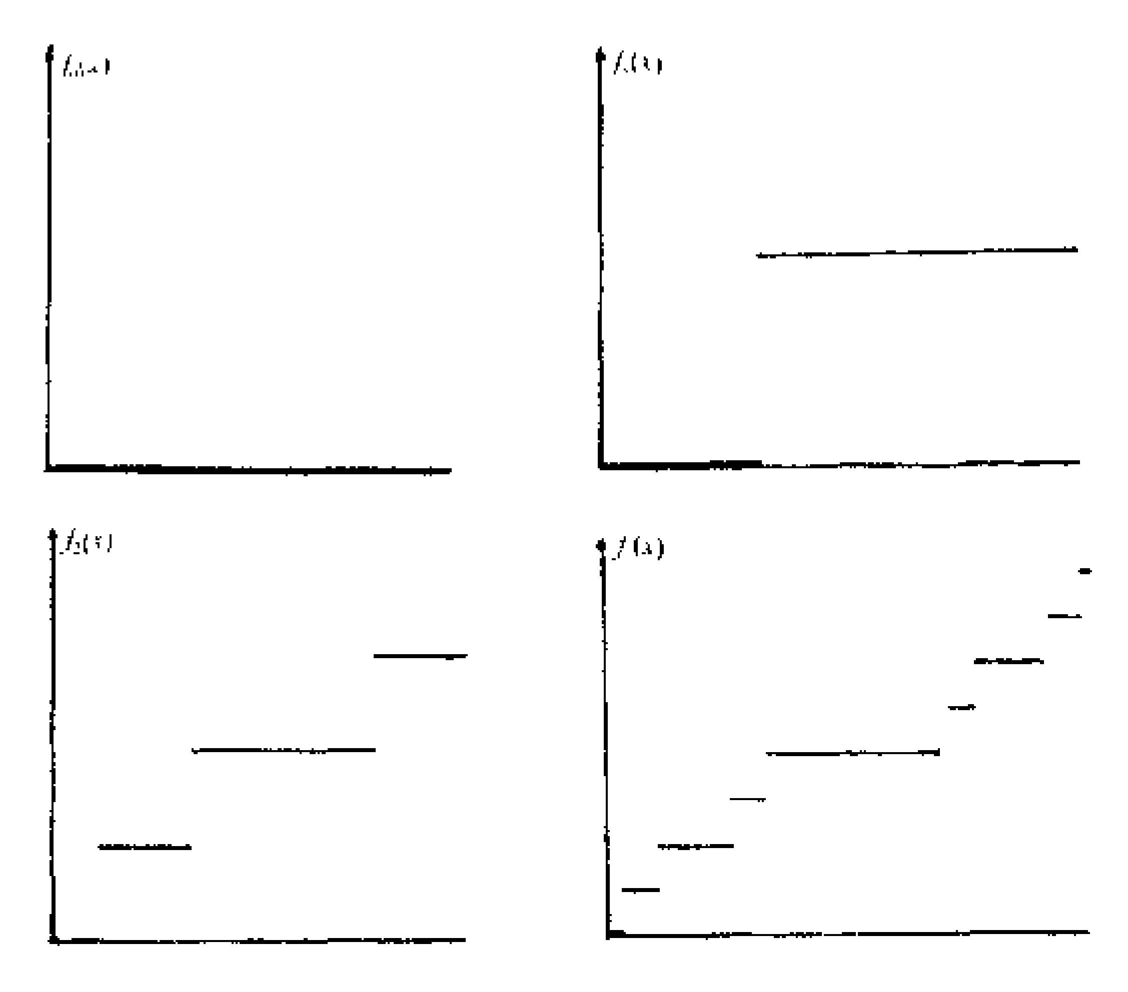


图3.5 f_0, f_1, f_2 和 f 的图示

§4 一维数据压缩与分形插值

1. 一维数据压缩

有一种平面图象的压缩方法，实行起来特别简单. 如对平面灰度图象采用§3 中介绍的象素数据模型，设一幅图象含有 $M\times N$ 个象素. 如将象素的二维位置坐标

$$(x_1,y_M),\quad (x_2,y_M),\quad \cdots\cdots,\quad (x_N,y_M),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$(x_1,y_1),\quad (x_2,y_1),\quad \cdots\cdots,\quad (x_N,y_1),$$

重新按某种顺序排为一维数组

$$u_1,u_2,\cdots\cdots,u_{N\cdot M}, \tag{4.1}$$

则该图象就对应于(4.1)式上的灰级，记作

$$z_1, z_2, \cdots\cdots\cdots, z_{NM}, \tag{4.2}$$

即 z_i 代表象素 u_i 的灰度($i=1,2,\cdots,NM$).

有两种方法可将图象通过扫描处理把二维数据转换为一维数组.常用的方法是光栅扫描法.图 4-1 显示了光栅扫描方法，图 4-2 是扫描得到的灰阶图.

这种由光栅扫描法而得到的图象编码(4.2)式，虽然方法简单，但由于当从一个区域穿过边界进入另一区域时，往往由于象素的灰度变化较大，造成灰级图曲线的剧烈抖动(如图 4-2 所示)，影响压缩比和解码后的图象质量.所以，更好的方法是第二种扫描法，即皮亚诺扫描法.这是利用分形曲线——皮亚诺曲线(§1 例 4)的性质，按块进行扫描，如图 4-3.它有两个重要特性：(1)原平面图象中的象素与扫描后一维序列中的元素是 1-1 对应的.(2)在一维序列中相邻的元素在原平面图象中一定也是相邻的象素.由于皮亚诺曲线的分维等于 2，它可以覆盖图象.而这种分块扫描的方法，使得边界上的象素以外，一个块内部的象素之间，灰度相差不大，可以减少灰阶曲线的剧烈抖动，因此可以提高图象编码的性能.

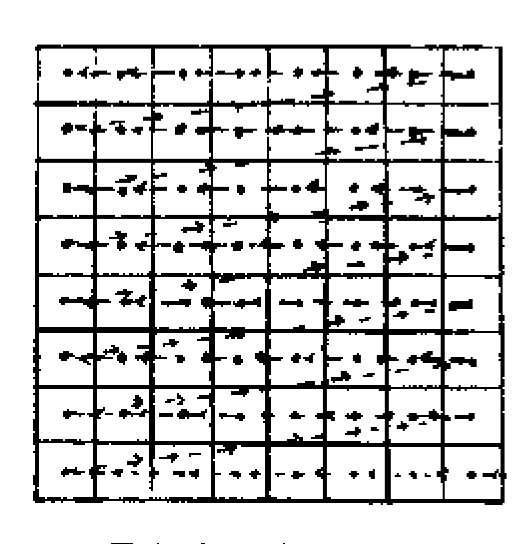


图 4-1 光栅扫描顺序

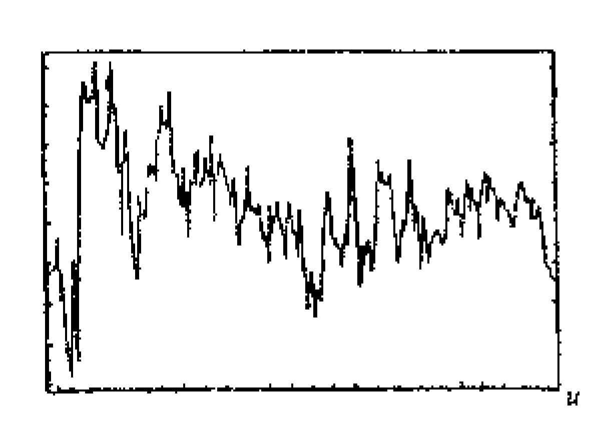


图 4-2 扫描后的灰阶图

得到了(4.2)式后，通常就可作压缩了.

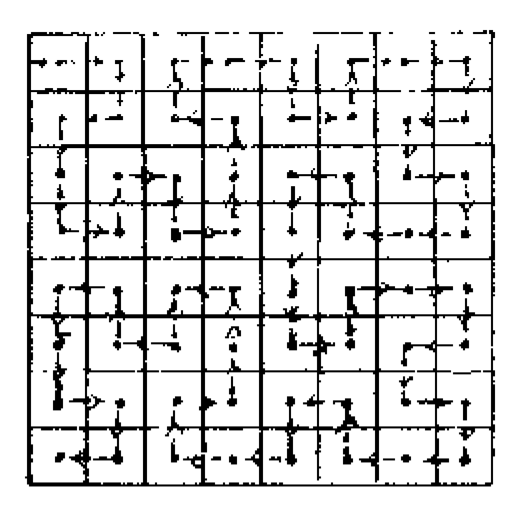


图 4-3 皮亚诺扫描顺序

不论由哪种方法扫描获得的一维灰级数据，(4.2)式都是精确的，无任何近似．由图 4-2 可见，灰级曲线图类似于海岸线那样弯弯曲曲，因此，为了压缩该数据，就像测量海岸线长度那样，选择某个长度为 l 的量尺，从边缘点 u_1 对应的 z 开始来度量灰级曲线．因为数组(4.1)式是离散的，因此每次都近似取量尺与曲线的交点是最靠近尺端的 z 点(当象素点之间距离充分小时就是合理的)，如图 4-4，图上已对量尺与曲线的交点坐标依次编号．令 $t_i=x_{i+1}-x_i$，则 $0<t_i\leqslant y$，及

$$g_{i+1}=g_i-\delta_i\sqrt{l^2-t_i^2}, \tag{4.3}$$

其中 g_i 是点 x_i 的灰级，δ_i 是符号函数，即

$$\delta_i=\begin{cases}1, & \text{当 } g_{i+1}>g_i \text{ 时；}\\ -1, & \text{当 } g_{i+1}<g_i \text{ 时．}\end{cases}$$

因此，如果取 $l>1$(设原始象素序列(4.1)式中各点距离为 1)，(t_i,δ_i)就是图象的有失真压缩编码．l 越大，压缩比也就越大．而由这组编码及起始点的灰级，就可恢复出原灰级曲线的近似曲线，如图 4-5．

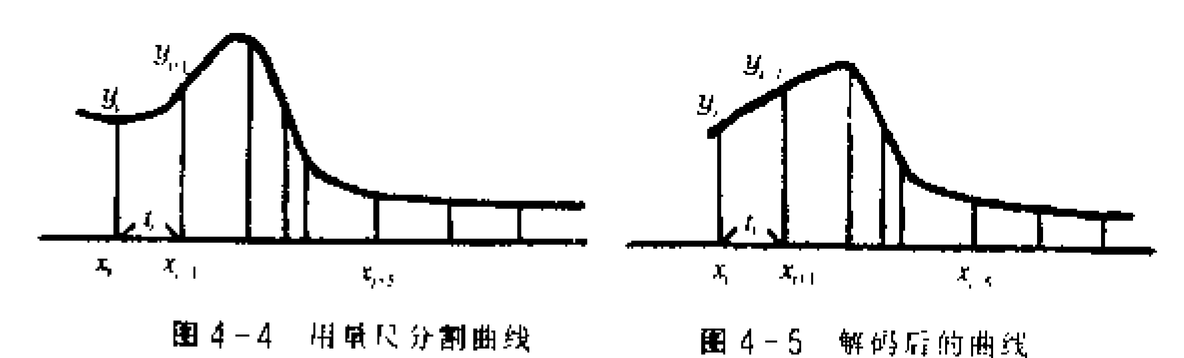


图 4-4 用量尺分割曲线　　图 4-5 解码后的曲线

2. IFS 方法构造的分形插值函数

上面介绍的压缩编码方法，实际上是线性插值函数．方法虽然

简单，但由于灰级曲线(如图 4－2 所示)的波动性很大，很不规则，线性插值的效果就不好. 能否利用分形曲线的特性引入图象的一维序列压缩呢？下面要介绍的分形插值函数，就是基于这样的假设：图象的灰级波形曲线是分形曲线，证明可见[16].

定义 1 设有一组数据，即点列

$$\{(x_i, F_i) \in \mathbb{R}^2 : i = 0,1,2,\cdots,n\}, \qquad (4.4)$$

其中

$$x_0 < x_1 < x_2 < \cdots < x_n.$$

对应于该组数据的**插值函数**就是连续函数 $f:[x_0,x_n]\to\mathbb{R}$，满足

$$f(x_i) = F_i, \qquad i = 1,2,\cdots,n \quad (4.5)$$

又，称点 $(x_i,F_i)\in\mathbb{R}^2$ 是**插值点**，并称函数 f 的图**通过**插值点.

设给出一组形如(4.4)的数据，现在要构造一个 IFS，使它的吸引子就是函数 f 的图(即通过插值点). 而且，像以前一样，为了计算方便，规定 IFS 中都是仿射变换. 设 IFS 为$\{\mathbb{R}^2; w_i, i=1,2,\cdots,n\}$，其中的映射 w_i 为

$$w_i\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a_i & 0 \\ c_i & d_i \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e_i \\ f_i \end{pmatrix}, \qquad i=1,2,\cdots,n \quad (4.6)$$

且满足

$$w_i\begin{pmatrix} x_0 \\ F_0 \end{pmatrix} = \begin{pmatrix} x_{i-1} \\ F_{i-1} \end{pmatrix}, \quad w_i\begin{pmatrix} x_n \\ F_n \end{pmatrix} = \begin{pmatrix} x_i \\ F_i \end{pmatrix}, \qquad i=1,2,\cdots,n \quad (4.7)$$

满足(4.7)式的仿射变换(4.6)的几何意义就是作“剪切变形”，即将(两边与 y 轴平行的)平行四边形 π 变形为同类平行四边形 $w_i(\pi)=\pi_i$，如图 4－6 所示. (4.6)式确定的仿射变换系数中已规定 $b_i=0$，所以还有 5 个参数 a_i, c_i, d_i, e_i 和 f_i 要确定，它们必须满足(4.7)的 4 个线性代数方程，所以有 1 个参数是自由的，我们选作 d_i，于是，根据给定的数据(4.4)式和 d_i，就可解出

$$a_i = \frac{x_i - x_{i-1}}{x_n - x_0}, \qquad (4.8)$$

$$e_i = \frac{x_n x_{i-1} - x_0 x_i}{x_n - x_0}. \qquad (4.9)$$

$$c_i=\frac{(F_i-F_{i-1})}{x_n-x_0}-d_i\frac{F_n-F_0}{x_n-x_0},\tag{4.10}$$

$$f_i=\frac{x_nF_{i-1}-x_0F_i}{x_n-x_0}-d_i\frac{x_nF_0-x_0F_n}{x_n-x_0}.\tag{4.11}$$

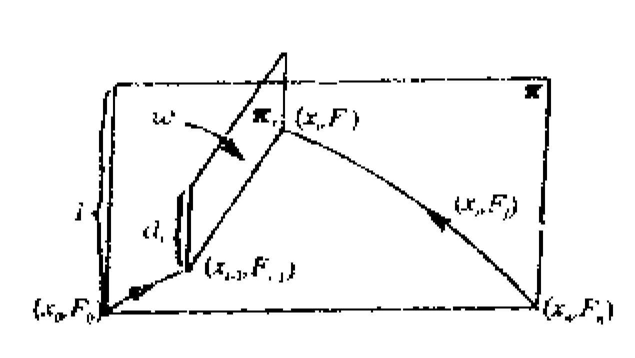


图 4-6 变换 w_i 将 π 剪切变形为 π_i

其中 $i=1,2,\cdots,n$. 我们把 d_i 称为变换 w_i 的**竖直标度因子**. 因为如取平行于 y 轴的直线段 L, 其长度为 1, 则经剪切变换 w_i 后仍为平行于 y 轴的直线段 $w_i(L)$, 其长度为 d_i(见图 4-6). 所以选定系数 d_i, 就是确定了由变换 w_i 所产生的竖直变换. 例如, 当取 $d_i=0$ $(i=1,2,\cdots,n)$, 就回到前面的线性插值函数.

现在我们已经可以从数据集(4.4)求出一组仿射变换(4.6). 与§2 的 IFS 方法一样, 我们要证明仿射变换的压缩性, 然后就可应用不动点定理和拼贴定理来实现图象压缩. 于是有下面的定理[3].

定理 1 设 n 是大于 1 的正整数, 记 $\{\mathbb{R}^2;w_i:i=1,2,\cdots,n\}$ 是对应于数据集 $\{(x_i,F_i):i=1,2,\cdots,n\}$ 的 IFS, w_i 的形式如(4.6)式. 又设竖直标度因子 d_i 满足 $0\leqslant d_i\leqslant 1(i=1,2,\cdots,n)$, 则存在一个 $\mathbb{R}^2$ 上的度量 d, 使得该 IFS 关于 d 是压缩的. 特别, 存在唯一的非空紧集 G, 满足

$$G=\bigcup_{i=1}^{n}w_i(G).\tag{4.12}$$

证明 由下式定义 $\mathbb{R}^2$ 上的度量 d:

$$d((x_1,y_1),(x_2,y_2))=|x_1-x_2|+\theta|y_1-y_2|.\tag{4.13}$$

其中 θ 是正实数, 将于后面来确定. 容易证明该度量与 $\mathbb{R}^2$ 上的欧氏度量是等价的. 现设 $n\in\{1,2,\cdots,n\}$, 而系数 a_n、c_n、e_n 和 f_n 由(4.8)~(4.11)式定义, 则有

$$d(w_i(x_1,y_1),w_i(x_2,y_2)) = d((a_ix_1+e_i,c_ix_1+d_iy_1+f_i),$$
$$(a_ix_2+e_i,c_ix_2+d_iy_2+f_i))$$
$$= a_i|x_1-x_2|+\theta|c_i(x_1-x_2)+d_i(y_1-y_2)|$$
$$\leqslant (|a_i+\theta|c_i|)|x_1-x_2|+\theta|d_i||y_1-y_2|.$$

因为$|a_i|=\frac{|x_i-x_{i-1}|}{|x_n-x_0|}<1$(因$n\geqslant 2$),所以选$\theta$为

$$\theta=\begin{cases}1, & \text{当 } c_i=0 \text{ 时;}\\ \frac{\min\{1-|a_i|:i=1,2,\cdots,n\}}{\max\{2|c_i|:i=1,2,\cdots,n\}}, & \text{其他,}\end{cases} \tag{4.14}$$

就有

$$d(w_i(x_1,y_1),w_i(x_2,y_2))$$
$$\leqslant(|a_i|+\theta|c_i|)|x_1-x_2|+|d_i||y_1-y_2|$$
$$\leqslant a|x_1-x_2|+\theta\delta|y_1-y_2|$$
$$\leqslant\max(a,\delta)d((x_1,y_1),(x_2,y_2)).$$

其中

$$a=\left(\frac{1}{2}+\frac{\max\{|a_i|;i=1,2,\cdots,n\}}{2}\right)<1,$$
$$\delta=\max\{|d_i|:i=1,2,\cdots,n\}.$$

定理2 设n是大于1的正整数.记$\{\mathbb{R}^2;w_i:i=1,2,\cdots,n\}$是对应于数据集$\{(x_i,F_i):i=1,2,\cdots,n\}$的IFS,$w_i$的形式如(4.6)式.又,设竖直标度因子$d_i$满足$0\leqslant d_i\leqslant 1(i=1,2,\cdots,n)$,使得IFS是压缩的.又记$G$为IFS的吸引子.则$G$就是连续函数$f:[x_0,x_n]\to\mathbb{R}$的图.而$f$是数据$\{(x_i,F_i):i=1,2,\cdots,n\}$的插值函数,即

$$G=\{(x,f(x)):x\in[x_0,x_n]\}, \tag{4.15}$$

其中

$$f(x_i)=F_i, \qquad i=0,1,2,\cdots,n \tag{4.16}$$

证明 记$\mathscr{F}$为满足下面条件的连续函数$f:[x_0,x_n]\to\mathbb{R}$的集合:$f(x_0)=F_0$和$f(x_n)=F_n$.仍取F上的尺度d是上界范数,即

$$d(f,g)=\max\{|f(x)-g(x)|:x\in[x_0,x_n]\},$$
$$\forall f,g\in\mathscr{F} \tag{4.17}$$

则(F,d)是完备度量空间[1].

设实数 a_i, c_i, e_i 和 f_i 由(4.8)~(4.11)式所定义. 现在定义一个新的映射 $T: F\to F$ 为

$$Tf(x) = c_i l_i^{-1}(x) + d_i f(l_i^{-1}(x)) + f_i,$$

$$x\in[x_{i-1},x_i],\quad i=1,2,\cdots,n \tag{4.18}$$

其中 $l_i:[x_0,x_n]\to[x_{i-1},x_i]$ 是可逆变换

$$l_i(x) = a_i x + e_i. \tag{4.19}$$

我们可以证明如上定义的 T 确是 F 到自身的映射. 设 $f\in F$, 则函数$(Tf)(x)$满足端点条件, 因为

$$\begin{aligned}(Tf)(x_0) &= c_1 l_1^{-1}(x_0) + d_1 f(l_1^{-1}(x_0)) + f_1\\ &= c_1 x_0 + d_1 f(x_0) + f_1\\ &= c_1 x_0 + d_1 F_0 + f_1 = F_0,\end{aligned}$$

而

$$\begin{aligned}(Tf)(x_n) &= c_n l_n^{-1}(x_n) + d_n f(l_n^{-1}(x_n)) + f_n\\ &= c_n x_n + d_n f(x_n) + f_n\\ &= c_n x_n + d_n F_n + f_n = F_n.\end{aligned}$$

我们也容易证明$(Tf)(x)$是区间$[x_{i-1},x_i]$$(i=1,2,\cdots,n)$上的连续函数, 因此$(Tf)(x)$在每一点 x_i 上$(i=1,2,\cdots,n-1)$都是连续的. 但在这些点上, 对每一点都可从两个不同区间上计算. 即对 $i\in\{1,2,\cdots,n-1\}$有

$$\begin{aligned}(Tf)(x_i) &= c_{i+1} l_{i+1}^{-1}(x_i) + d_{i+1} f(l_{i+1}^{-1}(x_i)) + f_{i+1}\\ &= c_{i+1} x_0 + d_{i+1} f(x_0) + f_{i+1} = F_i,\end{aligned}$$

当取 $x_i\in[x_i,x_{i+1}]$;

$$\begin{aligned}(Tf)(x_i) &= c_i l_i^{-1}(x_i) + d_i f(l_i^{-1}(x_i)) + f_i\\ &= c_i x_n + d_i f(x_n) + f_i = F_i.\end{aligned}$$

当取 $x_i\in[x_{i-1},x_i]$.

即两种计算导致相同的结果. 所以 T 确是从 F 映射到 F.

现在证明 T 是度量空间$(\mathscr{F},d)$上的压缩映射：设 f，$g\in\mathscr{F}$，$n\in\{1,2,\cdots,n\}$而 $x\in[x_{i-1},x_i]$，则

$$\begin{aligned}|(Tf)(x)-(Tg)(x)| &= |d_i|\,|f(l_i^{-1}(x))-g(l_i^{-1}(x))|\\ &\leqslant |d_i|\,d(f,g).\end{aligned}\tag{4.20}$$

现取 $\delta=\max\{|d_i|: i=1,2,\cdots,n\}<1$，则

$$d(Tf,Tg)\leqslant\delta d(f,g).\tag{4.21}$$

至此，我们就证明了 T：$\mathscr{F}\to\mathscr{F}$ 是压缩映射. 于是，根据压缩映射定理(§2)，T 在 $\mathscr{F}$ 中具有唯一的不动点. 即存在一个函数 $\bar{f}\in\mathscr{F}$，使得

$$(T\bar{f})(x)=\bar{f}(x).\qquad \forall\, x\in[x_0,x_n]\tag{4.22}$$

显然，$\bar{f}$ 是通过所有插值点的.

现在证明 $\bar{f}$ 的图就是 $\bar{G}$. 假设 $\bar{f}$ 的图是 G'，则因为 T 的定义式又可写成

$$(Tf)(a_ix+e_i)=c_ix+d_if(x)+f_i.$$
$$x\in[x_0,x_n],\quad i=1,2,\cdots,n$$

因此有

$$G'=\bigcup_{i=1}^{n}w_i(G').\tag{4.23}$$

而且 G' 是 $\mathbb{R}^2$ 的非空紧子集. 但由定理 1，只存在一个非空紧集 $\bar{G}$(即 IFS 的吸引子)，使(4.12)式成立，所以只能有

$$\bar{G}=G'.$$ □

定理 2 证明中的映射 T，即(4.18)式，将在图象解码中用到.

定义 2 对于满足上述定理 1 和定理 2 条件的 IFS，其吸引子是函数 $\bar{f}(x)$的图，则称 $\bar{f}(x)$为对应于数据集$\{(x_i,F_i): i=1,2,\cdots,n\}$的**分形插值函数**.

3. 用一维分形插值函数实现图象压缩

根据前述的有关理论，具体实现原理如下：

(1)首先将平面灰度图象 L 离数化，在各个象素(2 维坐标点)

上有相应的灰阶值，然后根据皮亚诺扫描法得到一组重新排列的一维序列$\{u_1, u_2, \cdots, u_m\}$（象素）及相应的灰阶值$\{z_1, z_2, \cdots, z_m\}$.

(2)根据对图象压缩比及清晰度的要求，选择适当的量尺的长度l（相邻象素间距离的整数倍）. 一般说来，l越大，压缩比就越高，但得到的图象质量就要差一些. 现在从u_1开始每隔l取一点，直到u_m，把所得到的点列重新记作$(x_0, x_1, x_2, \cdots, x_n)$，其中$x_0 = u_1$，$x_n = u_m$；同时，将相应的灰阶值重新记作$(F_0, F_1, F_2, \cdots, F_n)$，于是，我们就得到了对应于图象$L$的压缩数据集$\{(x_i, F_i); i = 0, 1, 2, \cdots, n\}$，也就是完成了编码.

(3)构造对应于数据集$\{(x_i, F_i); i = 0, 1, 2, \cdots, n\}$的IFS$\{\ ;w_i; i = 1, 2, \cdots, n\}$. 即先由(4.8)～(4.11)式求出仿射变换$w_i$的系数$a_i, e_i, c_i$和$f_i$，并选定适当的$d_i$，要求$0 \leqslant d_i < 1$. 于是就得到$n$个仿射变换$w_i$，如(4.6)式.

(4)构造新的映射T，然后通过迭代求得不动点，即函数$\bar{f}(x): [x_0, x_n] \to$. 先按(4.18)式构造T. 不妨取$f_0(x) = 0$，然后反复应用变换T，根据$f_{i+1}(x) = (Tf_i)(x)$求出函数序列$\{f_i; i = 0, 1, 2, \cdots, n\}$. 它的极限就是$\bar{f}(x)$，即对应于$\{(x_i, F_i)\}$的分形插值函数.

图4－7是柯西收敛序列$\{f_0(x), f_1(x), f_2(x), \cdots\}$的示意图.

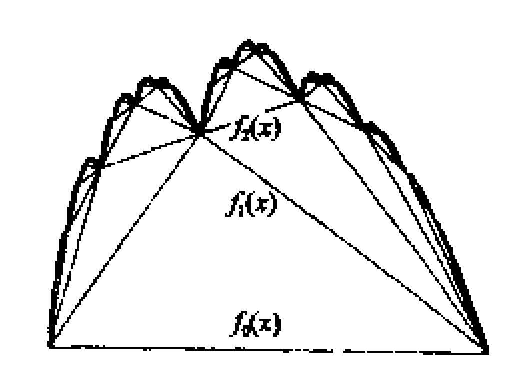


图4－7 收敛函数序列$\{f_{i+1}(x) = (Tf_i)(x); i = 0, 1, 2, \cdots\}$

以上介绍的是确定性算法. 当然，也可以用§3介绍的随机迭代算法，来求IFS的吸引子$\bar{f}(x)$.

(5)令$\bar{z}_i = \bar{f}(u_i)$ $(i = 1, 2, \cdots, m)$，再根据皮亚诺曲线的“反扫描”，就最后得到一幅离散化的灰度图象L'. 它是原图象L的近似.

以上一维分形插值方法具有简单

快速、存储量小等优点，对于清晰度要求不十分高的图象很适用。也有直接用二维分形插值方法的，见[17]。

4. **分形插值函数的分维**

下面的定理告诉我们分形插值函数的分维公式。因此，我们可以通过对竖直标度 d_i 的控制，来修正压缩后恢复的图象 L' 与原图象 L 的吻合程度。

定理 3[9] 设 n 是大于 1 的正整数。而 $\{(x_i,F_i)\in \mathbb{R}^2; i=1,2,\cdots,n\}$ 是数据集。又设 $\{\mathbb{R}^2; w_i; i=1,2,\cdots,n\}$ 是对应于该数据集的 IFS，其中 w_i 如(4.6)式所示。而竖直标度因子 d_i 满足 $0\leqslant d_i<1$，系数 a_i,c_i,e_i 和 f_i 由(4.8)～(4.11)式确定($i=1,2,\cdots,n$)。记 G 是 IFS 的吸引子，因此 G 是对应于该数据集的分形插值函数的图。如果

$$\sum_{i=1}^{n}|d_i|>1, \tag{4.24}$$

而且所有插值点 $\{(x_i,F_i); i=1,2,\cdots,n\}$ 不共线，则 G 的分维 D 是方程

$$\sum_{i=1}^{n}|d_i|a_i^{D-1}=1 \tag{4.25}$$

的唯一实数解；否则，G 的分维是 1.

定理 3 的证明见[18]。

[例 1] 设某分形插值函数的插值点是等间距的。现讨论它的分维。

设 $x_i=x_0+\frac{i}{n}(x_n-x_0)$（其中 $i=0,1,2,\cdots,n$），则代入(4.8)式后，即得

$$a_i=\frac{1}{n}. \qquad (\text{其中 } i=1,2,\cdots,n)$$

如果条件(4.24)式成立，则由定理 3，该插值函数的分维 D 满足(4.25)式，即有

$$\sum_{i=1}^{n}|d_i|\left(\frac{1}{n}\right)^{D-1}=\left(\frac{1}{n}\right)^{D-1}\sum_{i=1}^{n}|d_i|=1.$$

由此解出

$$D = 1 + \frac{\log\left(\sum_{i=1}^{n} |d_i|\right)}{\log n}. \tag{4.26}$$

此公式告诉我们两件事：首先，因为$|d_i|<1$，所以$\sum_{i=1}^{n}|d_i| < n$. 又因满足(4.24)式$\sum_{i=1}^{n}|d_i| > 1$，代入(4.26)式，即得

$$1 < D < 2.$$

即此类分形插值函数的分维D小于2，但可任意接近2；而且D又可逐渐变化几乎下降到1. 其次，从(4.26)式可知，分维D不依赖于数据集$\{(x_i, F_i)\}$中的值$\{(F_i: i=0,1,2,\cdots,n\}$. 因此，我们可以找到一组具有相同分维的分形插值函数，只要使竖直标度因子满足下式就可以了：

$$\sum_{i=1}^{n}|d_i| = n^{D-1}. \tag{4.27}$$

这个性质对于我们选择合适的IFS，提高压缩图象的质量十分有用.

第2章

基于分形理论的图象压缩方法

本书第1章讨论了分形理论、自仿射迭代函数系统，拼贴定理及分形图象压缩的基本原理，本章介绍分形图象压缩在有灰阶的静态图象（如照片，图画）和运动图象（电视、电影）中的应用. 为了使对图象压缩编码不甚熟悉的读者能理解，我们首先介绍图象压缩的一般知识，然后介绍针对图象压缩问题的几种分形压缩方法.

§5　图象压缩的基本原理和判别标准

现代通信技术的发展，提出了大量地存贮、记录和传输各种静态图象和运动图象的问题. 这里所说的静态图象，包括了各种图片，如气象云图、遥感图象、真迹等. 而运动图象，则有可视电话、工业电视、会议电视、彩色广播电视等活动图象. 高质量传输和存贮图象的一种途径，是把连续的图象信息离散化，形成数字化图象；然后，对数字图象进行压缩编码，使其数据量大大小于原始的图象的数播量；接着是对压缩后的图象进行存贮或传输. 在使用时，再将编码后的图象恢复. 对数字图象的压编编码的处理过程，称为图象压缩. 原始的图象信号在通信中称为信源，而图象压缩在图象信号通信系统中称为信源压编编码，

一个静态的有灰度的图象 A，是指定义在

$$I^2 = \{(x,y): 0 < x < L_x, 0 < y \leqslant L_y\}$$

上的 f 的图. 这里，$f:\mathbb{R}^2\to\mathbb{R}$，而 f 的取值范围为

$$0 \leqslant f(x,y) < a, \qquad \forall\, x,y \quad (5.1)$$

其中，a 是图象的灰度的最大值，L_x、L_y 分别是图象的宽和高. 对二维空间 x、y 进行抽样（离散化），就可得到一个离散化的图象，

抽样后每个象点的亮度为$f(i,j)$，$i=1,2,\cdots N$，$j=1,2,\cdots M$. 其中N、M分别为x、y方向上抽样的点数. 我们称每个抽样点为一个象素(pixel). 这样，二维空间上的连续图象变成了$N\times M$个象素组成的离散图象. 其灰度也可离散为k个等级，此时f的图中的全体三元点$(x,y,f(x,y))$，变成了$(i,j,f_n(i,j))$，其中f_n为灰度的离散值，称为灰阶，$0\leqslant f_n<k$. 为了满足人肉眼看不出失真，一般取$k=256$. 图象的二维象素$N\times M$，可根据不同的精度要求分为352×288，720×480，720×576，512×512，1280×720个象素点.

活动图象$A_1,A_2,\cdots\cdots$则是一个随时间变化的图象组，可用给定的函数$g:\mathbb{R}^3\to\mathbb{R}$来表示，这里$g$就是指由全体四元点$(x,y,t,f(x,y,t))$所组成的$\mathbb{R}^4$中的一个子集. 对于离散化的数字图象组，其灰度函散$f(i,j,t)$被量化为k个等级，t也被离散化，即为帧数. 按照视觉能接收的标准，一般取25帧/秒或30帧/秒. 对于彩色图象，其信息量比黑白的图象增加3倍(红、黄、蓝). 例如，一幅512×512象素的灰阶图象，每个象素点具有256个灰阶($k=256$)，用8比特(bit)二进位表示，它将占有2兆比特(2 Mbits)存储空间；彩色情况下，则要占6兆比特存储空间. 如以每秒25帧计算，活动图象每秒的信息量为150兆比特. 这种"海量"信息，如果不进行压缩，则图象的存储、传输都会发生很大的困难.

由于一幅图象的各象素之间存在着很大的相关性，这就可利用一些编码的方法删去相关性，达到减少冗余的目的. 图象的冗余包括空间冗余(象素点之间存在的相关性)、时间冗余(活动图象两个连续帧之间的冗余)、信息熵冗余(单位数据量大于其熵)、结构冗余(区域上存在非常强的纹理结构)、知识冗余(有固定的结构如人的头像)、视觉冗余(有些图象的失真是人主观感觉不到的). 这些冗余可以采用各种不同的方法删去，而使信息达到压缩目的. 压缩的系统如图5－1所示. 原始图象组$A_1,A_2,\cdots$经过图象压缩处理后形成$B_1,B_2,\cdots$，然后经存储媒体(磁盘，光盘)存储成文件. 在使用时，解压缩获得恢复图象$A'_1,A'_2,\cdots$. 或者$B_1,B_2,\cdots$的图象

组经过通信媒体进行信道的编码和解码得到 $B'_1, B'_2, \cdots$ 的图象组，再经解压缩获得恢复图象组 $A'_1, A'_2, \cdots$. 图 5－1(i)、(ii)表示这两种不同的情况.

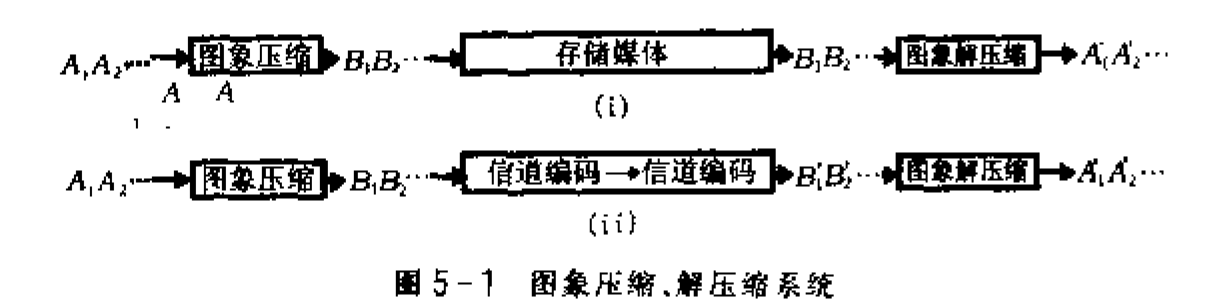


图 5－1 图象压缩、解压缩系统

§5.1 图象压缩编码的分类

图象压缩编码可分为无失真编码(或可逆压缩)与限失真编码(不可逆压缩)两大类. 从图 5－1 看，原始图象 A 经图象压缩到 B，然后经存储媒体或传输媒体(信道编码解码)得到 B'，再经信源解压缩获得恢复图象 A'，如 $A'=A$，则为可逆压缩；如 $A'\approx A$，为不可逆限失真压缩.

1. 无失真编码

无失真编码是将输入的图象中表达象素点灰阶的值的每个符号，用规定的码字符号按一定的方式编排而成. 由于规定的码字符比原图象中的符号短，从而可用比较少的比特数来表达原图象的符号，达到图象压缩的目的. 在恢复图象时，只要把码字符与象素点的灰阶符号对应起来，就可无失真地恢复图象. 例如一个静态图象 A，由 512×512 个象素组成，每个象素上的灰阶用 8 比特字符表示，即有 256 种可能的符号；如果我们分析每个符号出现的概率，把概率出现大的符号用比 8 比特短的码字符表示(如 2 比特)，而概率出现少的符号用较长的码字符表示(如 10 比特)，这样，平均每个码字符长度(比特数)就比 8 比特小. 无失真编码的常用方法有霍夫曼编码、游程编码、算术编码等.[18]

按照仙农(C. E. Shannon)的理论，无论规定的码字符如何编排，其平均的码字符表达的比特数不可能小于原将号表达的图象

的熵值.熵值被定义为

定义 1 设信息源 $\mathbb{S}$ 的符号集为 $\{\mathbb{S}_1,\mathbb{S}_2,\cdots,\mathbb{S}_N\}$，$\mathbb{S}_i$ 出现的概率为 $p(\mathbb{S}_i)$，则信息源的熵 $H(\mathbb{S})$ 为

$$H(\mathbb{S}) = -\sum_{i=1}^{N} p(\mathbb{S}_i)\log_2 p(\mathbb{S}_i). \tag{5.2}$$

$H(\mathbb{S})$ 的单位为：比特数/字符.

由于 $0\leqslant p(\mathbb{S}_i)\leqslant 1$，故有

$$H(\mathbb{S}) \geqslant 0.$$

当信源每个符号具有等概率分布时，即 $p(\mathbb{S}_i)=\frac{1}{N}, i=1,2,\cdots,N$，则有 $H_{\max}(\mathbb{S})=\log_2 N$，如果符号集中有一个 $P(\mathbb{S}_i)=1$，则有 $H_{\min}(\mathbb{S})=0$. 在无干扰情况下，存在一种无失真编码的压缩比下限，编码后每个字符的平均码长 L 与信源的熵 $H(\mathbb{S})$ 任意接近：

$$L = H(\mathbb{S}) + \varepsilon,$$

其中 ε 为任意小的正数.

原始图象的总比特数 N_0 和压缩后的图象总比特数 N_c 之比，称为图象的压编比. 即 $r=\frac{N_0}{N_c}$. 压缩比 r 也可用字符的平均码长来表示. 设 L_0 为原始图象数据的平均码长，无失真编码的最大压缩比为

$$r_{\max} = \frac{L_0}{H(\mathbb{S})}. \tag{5.3}$$

十分明显，无失真压缩不能取得高的压缩比，这是因为它受到信息源本身的熵的限制. 因而，无失真编码又称为熵编码.

2. 限失真编码

为了进一步提高图象编码的压缩比，利用图象中象素之间的相关性，以及人的视觉对图象灰度灵敏度的差异，限失真编码便成了压缩图象的重要手段. 常用的方法是：

i. 变换编码. 把原始图象通过一些变换后，得到在变换域中系数具有比较低的熵值，然后再用无失真的编码把信息压缩下去. 如

离散余弦变换(DCT)，小波变换(WT)，K－L 变换等.

ii. 预测编码. 利用相邻象素之间存在的相关性,采用预测值与原象素的灰阶值之差进行编码. 由于误差信号的熵值比较低,这时再用无失真编码就可进一步压编图象信息.

iii. 矢量量化. 把图象分块,每一个块的象素组成一个矢量,对于矢量进行新的编码. 注意到在图象中的很多块可能比较相近,因而只要用一个矢量就可代表这些块. 通过聚类算法,可用比较少的码把图象恢复出来.

iv. 模型法. 把图象分割成几个基本模型,模型具有一定的参数,只要将参数进行编码传输,就可恢复出原来的图象.

基于分形理论的图象压缩方法介于 iii、iv 之间,但又有其独特的一面,即把任何一幅图象看作一个分形,只要找到迭代函数的参数,那么就可用迭代的方法把原来的图象恢复出来. 图象既是被近似看作是分形,因面它是限失真的编码.

实际图象压缩编码系统中,是将无失真编码和限失真编码结合在一起使用,各种编码方法又可以交叉使用. 用于通信媒体的系统,图象编码后的信息以每秒一定的比特数传输(如 64 kbit/秒，1.5 Mbit/秒),以适应于通信编路的需要. 每秒传输的比特数又称为散码率或比特率. 对于高比特率的传输通道,其图象的压缩比可低一些,但对低比特率的传输通道(如电话线),需要对图象进一步的压缩,才能满足传输通道的要求.

§5.2 现有的数字图象编码方法

现有的图象压缩方法,都建立在离散余弦变换的基础上. 为了减少计算量,往往把一幅图象划为 8×8 象素点的子块,对于每个子块进行二编离散的余弦变换. 以$(x(m,n))_{8\times 8}$表示一个 8 行、8 列的数字图象子块矩阵,二维的离散余弦变换是将图象子块从空间域(m,n 平面)转换到频率域. 子块的二维余弦变换为

$$Y(k,l) = \frac{1}{4}c(k)c(l)\sum_{m=0}^{7}\sum_{n=0}^{7}x(m,n) \cdot \cos\frac{2m+1}{2\times 8}k\pi\cos\frac{2n+1}{2\times 8}l\pi, \qquad (5.4)$$

其中 $k=0,1,\cdots,7$；$l=0,1,\cdots,7$；

$$c(k)=\begin{cases}1/\sqrt{2}, & \text{当 } k=0 \text{ 时；}\\ 1, & \text{当 } k=1,2,\cdots,7 \text{ 时}.\end{cases}$$

$$c(l)=\begin{cases}1/\sqrt{2}, & \text{当 } l=0 \text{ 时；}\\ 1, & \text{当 } l=1,2,\cdots,7 \text{ 时}.\end{cases}$$

$Y(0,0)$为最低频率，$Y(0,1)$、$Y(1,0)$为次低频率，把 $Y(k,l)$按照从低频到高频系数排列，对于每个图象块，能量往往集中在低频，而高频分量很小. 采用对人眼最佳效果的加权函数来量化 $Y(k,l)$. 对低频分量，量化步长小；对高频分量，量化步长大. 小于量化级的，系数为 0. 通过量化，丢弃对图象解码后的质量影响不大的那些系数. 此时，每个 8×8 象素的数字图象块，它量化后的余弦变换系数有大量的"0"，这就很容易采用熵编码的方法，把信息压缩下来. 在使用时，解码的方式是可逆的，即熵解码，逆量化，逆余弦变换，就可恢复图象.

对于静止图象，余弦变换方法的压缩效果很好. 国际上，1990年制定了用于会议电视的 H. 261 标准，以及1991年由联合象片专家组(Joint Photographic Expert Group)制定的静态图象压缩标准，均采用了分块的余弦变换方法. 1991 年制定的标准称为JPEG，它是采用一定的头文件说明，加上分块余弦变换，量化，熵编码组成. 制定该标准是为了在国际范围内图象压缩的产品或国际间的交往有一个共同的语言. 按照 JPEG 标准，一幅静态图象 A，需要经图 5-2(i)的过程，才能得到压缩编码的目的；而重建图象，又要通过图 5-2(ii)的过程进行恢复，其核心是离数的余弦变换.

对于活动图象，压编方法也建立在分块的离散余弦变换的基

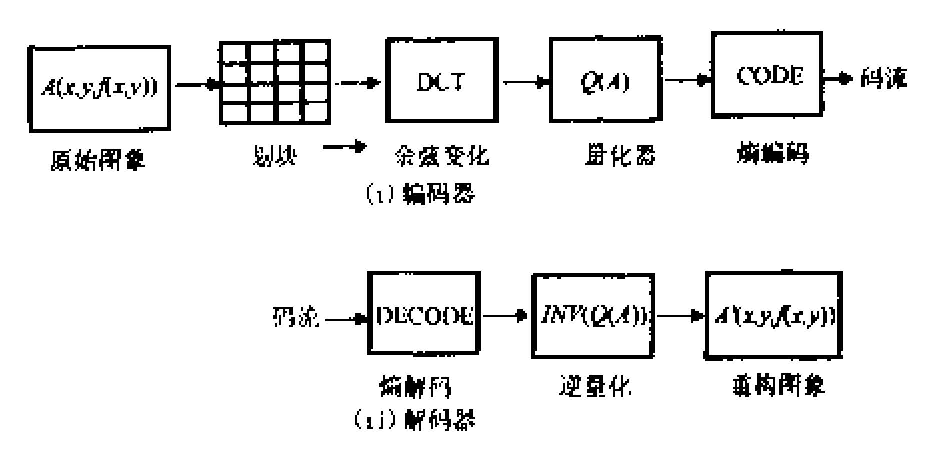


图5-2 JPEG标准的构成

础上. 对一个孤立的与前后帧不相关的图象，仍用静止图象的编码方式(称为帧内编码). 对于与前后帧有密切相关的帧，可用帧间预测的方法进行编码. 对没有运动的块，就可用前一帧的块来替代. 对于运动的块，采用运动矢量估计和两帧间的误差值进行编码. 由于大多数情况下，活动图象相邻两帧间的变化很小，因而误差值也很小，对于误差值进行余弦变换和熵编码，就可达到很高的压缩比.

为了对活动图象编码的数据有一个统一的格式，1991 年 3 月由运动图片专家小组(Motion Picture Experls Group)制定了 MPEG 1 的标准. 它适合于低分辨率的图象(352×240象素或352×288象素)，它的数码率为 1.5Mbit/秒. 1993 年 10 月，这个专家小组又制定了 MPEG 2 标准. 这标准适合于高分辨率(704×480象素或704×576象素)，它的数码率为 4～8 Mbit/秒，特别适宜于高质量的图象和声音信号，用于高清晰数字电视，高质量的电影光盘等. 更低码率的标准如 MPEG 4 和 MPEG 7 也将推出.

§5.3 图象质量的判别标准

对于限失真编码，原码 A 与重建图象 A' 之间存在着差异. 差异的大小，意味着恢复图象的质量不相同. 但是，由于人的视觉冗余度，对有些差异灵敏度比较差，这就产生两种判别标准；一种是客观判别标准，它建立在 A 和 A' 之两的误差上；另一种是主观评

价，用人的肉眼对图象打分而得到.

1. 客观评价

设原始的二维灰阶图象 $A=f(i,j)$，其中 $i=1,2,\cdots,N$；$j=1,2,\cdots,M$，经压缩还原后的图象数据为 $A'=f'(i,j)$，$i=1,2,\cdots,N$；$j=1,2,\cdots,M$. 可用下列几种指标进行评价：

均方误差 $$\mathrm{MSE}=\frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}(f(i,j)-f'(i,j))^2; \tag{5.5}$$

规范化均方误差 $$\mathrm{NMSE}=\frac{\mathrm{MSE}}{\sigma_f^2}, \tag{5.6}$$

其中 $\sigma_f^2=\frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M}f^2(i,j)$；

对数信噪比 $$\mathrm{SNR}=10\log\frac{\sigma_f^2}{\mathrm{MSE}}=-10\log\mathrm{NMSE}(\mathrm{db}); \tag{5.7}$$

峰峰对数信噪比 $$\mathrm{PSNR}=10\log\frac{255^2}{\mathrm{MSE}}(\mathrm{db}). \tag{5.8}$$

这里，db 为对数信噪比的单位(分贝).

十分明显，以上评价完全取决于原始图象每个象素上灰阶值与重建图象的误差. 这种评价在主观感觉上也有一定的参考意义. 实际上，常用的指标为 PSNR. 当 PSNR 超过 30 db 时，人的主观感觉很难找出其差异.

2. 主观评价

采用平均判分 MOS(mean opinion score)或多维计分等方法进行测试. 即组织一群足够多的实验人员，通过观察来评定图象的质量. 观察者给评定的图象打上一定的质量等级，比较损伤程度，及与原图进行比较等方法，根据不同的质量打上 5 级，6 级或 7 级的评分制，最后用平均的办法得到图象的分数. 这样的评分虽然很花时间，但比较符合实际. 目前，国际的标准都是采用主观评价.

主观评价与客观评价之间有一定的联系，但不能完全等同. 由于客观评价比较方便，很有说服力，故在论文中常被采用. 但主观

评价很直观，符合人的视觉效果，比较实际，在制定国际标准时常采用主观评价.

§6 分形静态数字图象压缩方法

一个确定性的严格自相似的分形图象$\overline{A}$，是完备度量空间$(\mathscr{H}(X),h)$中的一个子集，$\overline{A}\in\mathscr{H}(X)$. 存在一个 IFS，$\{X,w_1,w_2,\cdots,w_n,c\}$，$c<1$，$W(\overline{A})=\bigcup_{i=1}^{N}\hat{w}_i(\overline{A})$，$W:\mathscr{H}(X)\to\mathscr{H}(X)$为压缩变换，且有$W(\overline{A})=\overline{A}$. 图象的分形压缩，是给定一个$\overline{A}$，搜索一个 IFS 是压缩的，满足$W(\overline{A})=\overline{A}$；而图象的解压缩，则是已知 IFS，通过迭代函数系统，在任意初值$B\in\mathscr{H}(X)$下反复迭代，$\overline{A}=\lim_{n\to\infty}W^n(B)$，此极限图形就是原分形图象$\overline{A}$. 图象的编码就是量化$w_1,w_2,\cdots,w_N$的系数，并进行熵编码；而图象的解码则是熵解码，逆量化以及迭代. 如果N越小，则压缩比就越高. 例如§2 例 3 的枫叶，$N=4$，只要 4 组仿射压缩变换的系数就可完全恢复原始的图象，则量化后的系数为 24 个. 如原始图象是 512×512 个象素点的二值图象，每个象素点用 1 比特表示(黑，白)，而量化的参数用 8 比特表示，则枫叶图象的压缩比为 1365 倍，比目前现有的任何图象压缩方法的压缩比要高得多. 采用离散余弦变换为核心的方法，其压缩比在 8～20 倍左右；如通过精心调整的小波变换方法，也只能到 100 倍左右，可见分形压缩编码在高压缩比的图象压缩上，有它的巨大潜力.

现实生活中的图象是有灰度的、非严格自相似的图象，不具有整体和局部的自相似性. 严格分形图象在现实生活中只占极小的部分. 巴恩斯利[19]提出 IFSP，采用由初始点经压缩变换的参数$w_i(a_i,b_i,c_i,d_i,e_i,f_i)$加上伴随概率$p_i$进行迭代，得到的点集的分布类似于灰阶的效应，但在编码方面却未给出任何具体的算法. 一个大致的步骤为：把原图预分解为若干个分形的子图，使得每个子

图具有一定的分形结构(子图的整体与局部之间存在某种自仿射特征),分解可采用其他的图象处理手段,且由大量的这些子图组成了分形库,每个子图可在这些分形库中找到它们的匹配子图编码. 这样,对图象的分形编码可化为图象分割,到库中找匹配子图的编码,最后扔掉原图,保存子图编码,进行存贮或传输.

从目前来看,巴恩斯利的方法存在很大的问题. 因为对于图象的子图分割本身,还没有一种计算机自动分割的办法;其次,分形库的规模大,没有统一的建库的办法;高压缩比编码搜索的过程是以费时费力为代价,难以实现自动压缩. 有效的方法至少目前没有公开.

简库恩(A. E. Jacqain)[20],[21]提出了全自动的分形图象压缩方法. 该方法是以局部的仿射变换代替全局的仿射变换,基于图象划块的方式. 以后,弗歇尔[22],[23]改进了这个方法,此即目前分形图象编码中的主要方法.

设 A 为原始的一个有灰度的图象,用三元点的全体来表示 $(x,y,f(x,y))$. 其中函数 $f:\mathbb{R}^2\to\mathbb{R}$ 是二维平面上的灰度函数,$f(x,y)$的定义域为 $0\leqslant x\leqslant L_x$, $0\leqslant y\leqslant L_y$. $\{w_i\}$是 $\mathbb{R}^2$上的压缩变换族(其中 $i=1,2,\cdots,N$),w_i 的压缩因子为 $c_i<1$. $\varphi=\{\varphi_i$,其中 $i=1,2,\cdots,N\}$是对应每个 w_i 的灰度映射族,$\varphi_i:\mathbb{R}^+\to\mathbb{R}^+$. 图象 A 在(x,y)上的灰度,都可为 N 个 w_i 变换所对应的灰度变换后的累加. 变换用 T 表示,可写为

$$(Tf)(x,y)=\sum_{i=1}^{N}{}'\varphi_i(f(w_i^{-1}(x,y))). \tag{6.1}$$

上式中,w_i^{-1} 是仿射变换 w_i 的逆变换,$\sum'$ 代表灰度的累加只对 $w_i^{-1}(x,y)$在定义域范围的那些灰度值有意义,如果 $w_i^{-1}(x,y)$已在定义域外,就不需要累加.

从式(6.1)看到,带灰度的图象包含两种仿射变换,一种是在 $\mathbb{R}^2$上的压缩变换 $w_i,i=1,2\cdots N$,它是将图象 A 整个二维定义域变换到一些小的子块,再覆盖到整个定义域上. 另一种是在 $\mathbb{R}^+$上的

灰度变换 φ_i，它将图象灰度函数 $f(x,y)$ 通过线性变换与一些相应子块的灰度函数匹配.

设 A_i 是 A 上的一个子块，其灰度函数可表示为 $f(A_i)$. 在 $\mathbb{R}^2$ 上的压缩变换，是寻找一个 $\hat{w}_i$，使 $\hat{w}_i(A)=A_i$. 在 $\mathbb{R}^+$ 上的灰度变换 φ_i，是找到一个匹配系数 C_i 和 h_i，使变换 φ_i 前后灰度函数

$$f'(A_i) = C_i f_{av}(\hat{w}_i^{-1}(A_i)) + h_i \qquad (6.2)$$

和 $f(A_i)$ 之间的距离

$$d_p = \|f'(A_i) - f(A_i)\|_p \qquad (6.3)$$

达到最小. 上式中，$f_{av}(\hat{w}_i^{-1}(A_i))$ 表示在 $\hat{w}_i$ 映射前的灰度经多象素点的平均缩小到与子块 A_i 定义范围相同的尺度的函数. 在平方可积空间，d_p 为

$$d_p(f',f) = \left[\iint_R |f'(A_i) - f(A_i)|^p dxdy\right]^{\frac{1}{p}}. \qquad (6.4)$$

当 $p=2$ 时，对于离散化了的数字图象，积分号改为累加号，则 $f'(A_i)$ 与 $f(A_i)$ 之间的距离可表为

$$d_p(f'f) = \left[\sum_{i=1}^{M}\sum_{j=1}^{N}(f'(i,j) - f(i,j))^2\right]^{\frac{1}{2}}. \qquad (6.5)$$

上式中，M、N 分别为图象在 x、y 方向上的象素点数. 对照(6.5)与(5.5)式，d_p 的度量与数字图象在压缩后恢复的图象的客观评价是一致的，此时灰度也被离散为灰阶.

多数带灰阶的图象并不存在整体的自相似性，只能利用局部的自相似性来求出仿射变换系数，达到压缩编码目的. 下面介绍全自动局部迭代函数方法对有灰阶的数字图象编码方法. 不失一般性，我们讨论正方形的数字图象 A，它有 $2^N \times 2^N$ 的象素点，其 $(f(i,j),i,j)$ 被量化为256个灰阶 $f_k(i,j)$，$0 \leqslant k \leqslant 255$.

把图象分割成 n 个子块 R_i，$i=1,2,\cdots,n$，它的尺寸为 $2^R \times 2^R$，i 代表子块的编号：有 $R_i \cap R_j = \varnothing$，$i \neq j$；$\bigcup_{i=1}^{n} R_i = A$. 让 D_j 代表比 R_i 大的尺寸的图象子块，$2^D \times 2^D$，$D > R$，$D_i \cap D_j \neq \varnothing$，$j=1,2$，

$\cdots,(2^N-2^D+1)\times(2^N-2^D+1)$. 自相似性的匹配在 R_i 与 D_j 之间进行. 由于 D_j、R_i 比整个图象小得多，只要子块划得足够的小，局部的自相似性在图象中总是存在的. 按§3.3的定义，R_i 为子块，D_j 为父块.

为了找出两个仿射变换的系数之间的关系，现定义一些操作符如下：

"取块"操作符 $B^J_{K,L}$　它作用于图象 A 后表示在有灰阶的图上取出一个三维的块，其底的大小为 $2^J\times 2^J$，其第三维，即灰阶函数，取块连同 $2^J\times 2^J$ 的底一起取出来. K,L 表示子块在二维底部左上角的位置，坐标为 (K,L).

"放块"操作符 $(B^J_{K,L})^*$　$(B^J_{K,L})^*A$ 表示把底部 $2^J\times 2^J$ 的三维块插入全 0 的图中，且插入子块的左上角位于 (K,L) 上. 把块坐标(块左上角坐标)用大写字母表示，在块内的象素坐标用小写字母 (k,l) 表示.

"平均-抽样"操作符 A_v　把 $2^J\times 2^J$ 带有灰阶的图象通过近邻四个象素点灰阶求平均-抽样，变成 $2^{J-1}\times 2^{J-1}$ 尺寸的块. "A_v"操作可用下式表示：

$$(A_vB^J_{K,L}A)(k,l)=\frac{1}{4}[(B^J_{K,L}A)(2k,2l)+(B^J_{K,L}A)(2k+1,2l)$$
$$+(B^J_{K,L}A)(2k,2l+1)+(B^J_{K,L}A)(2k+1,2l+1)]. \tag{6.6}$$

通过 A_v 操作，$2^J\times 2^J$ 的图变为 $2^{J-1}\times 2^{J-1}$ 的图，其灰阶函数的分辨率也减小. 即变得粗糙了.

"反射-旋转"操作符 L_p　是表示块的旋转变换，犹如 IFS 变换中的系数 a,b,c,d. 为了简化匹配，把一个正方形子块的旋转化为最简单的8种方式，即旋转 0°, 90°, 180°, 270°和垂直中线反射、水平中线反射和对角反射这 8 种固定方式，$p=8$. 设 D_j 是一个图象块，它的灰阶函数为 $\{u_j(i_x,i_y):i_x,i_y=0,1\cdots,2^D-1\}$ 对应的八种变换为

L_1：　$u_j(i_x,i_y)=u_j(i_x,i_y)$，　　　　旋转 0°；

L_2：　$u_j(i_x,i_y)=u_j(i_x,2^D-1-i_y)$，　　　　垂直中线反射；

L_3：　$u_j(i_x,i_y)=u_j(2^D-1-ix,i_y)$，　　　　水平中线反射；

L_4：　$u_j(i_x,i_y)=u_j(i_y,i_x)$，　　　　对角线 $i_x=i_y$ 反射；

L_5：　$u_j(i_x,i_y)=u_j(i_y,2^D-1-i_x)$，　　　　对角线 $i_x+i_y=2^D-1$反射；

L_6：　$u_j(i_x,i_y)=u_j(i_y,2^D-1-i_x)$，　　　　旋转 90°

L_7：　$u_j(i_x,i_y)=u_j(2^D-1-i_x,2^D-1-i_y)$，　旋转 180°

L_8：　$u_j(i_x,i_y)=u_j(2^D-1-i_y,i_x)$，　　　　旋转 270°

利用上面定义的操作符，块与块之间的自相似性的搜索是去找一个近似的匹配，使

$$(B^R_{K,L}A) \approx C_{K,L}L_{P(K,L)}A_vB^D_{\Pi(K,L)}A + h_{K,L}B^R_{K,L}I. \qquad (6.7)$$

式中 $C_{K,L}$是位置在(K,L)子块的变换伸缩因子，$h_{K,L}$是它的偏移量，I 是 $2^N\times 2^N$ 的全 1 矩阵，$\Pi(K,L)$表示从位子(K,L)位置上的子块映射到对应的父块位置，$P(K,L)$表示对位子(K,L)上的子块的8种变换，$p=1,2,\cdots,8$.

编码的过程是：

1. 将图象 A 分割为互不交叠的子块 R_i，每个 R_i 块可用“取块”得到，符号为 $B^R_{K,L}A$.

2. 搜索与其自相似的父块 D_j，并做缩小，旋转和平移，父块尺寸 $2^D\times 2^D>2^R\times 2^R$，父块的取块变换符号为$(L_{P(K,L)}A_vB^D_{\Pi(K,L)}A)$，通过 A_v 后父块与子块有相同的尺寸.

3. 如果 $B^R_{K,L}A$ 和$(L_{P(K,L)}A_vB^D_{\Pi(K,L)}A)$各象素点上的灰阶值分别为 $a_1,a_2,\cdots,a_m$ 和 $b_1,b_2,\cdots,b_m$，则(6,7)近似式两边近似的误差可用(6,5)式，去掉开方，加入归一化因子 m 而得到：

$$d_P = \frac{1}{m}\sum_{i=1}^{m}(C_{K,L}b_i + h_{K,L} - a_i)^2. \qquad (6.8)$$

搜索是改变 $\Pi(K,L)$和 $L_{P(K,L)}$，使 d_p 达到最小. 即利用对(6.8)式

求导而得系数 $C_{K,L}$ 和 $h_{K,L}$. 令 $\frac{\partial d_p}{\partial C_{K,L}}=0$ 和 $\frac{\partial d_p}{\partial h_{K,L}}=0$, 得

$$C_{K,L}=\frac{m\left(\sum_{i=1}^{m}a_ib_i\right)-\left(\sum_{i=1}^{m}a_i\right)\left(\sum_{i=1}^{m}b_i\right)}{m\left(\sum_{i=1}^{m}b_i^2\right)-\left(\sum_{i=1}^{m}b_i\right)^2}. \tag{6.9}$$

且当 $m\left(\sum_{i=1}^{m}b_i^2\right)-\left(\sum_{i=1}^{m}b_i\right)^2=0$ 时, $C_{K,L}=0$,

$$h_{K,L}=\frac{1}{m}\left(\sum_{i=1}^{m}a_i-C_{K,L}\sum_{i=1}^{m}b_i\right). \tag{6.10}$$

此时, 有

$$d_p=\frac{1}{m}\Big[\sum_{i=1}^{m}a_i^2+C_{K,L}\Big(C_{K,L}\sum_{i=1}^{m}b_i^2-2\sum_{i=1}^{m}a_ib_i+2h_{K,L}\sum_{i=1}^{m}b_i\Big)$$
$$+h_{K,L}(h_{K,L}m-2\sum_{i=1}^{m}a_i)\Big]. \tag{6.11}$$

4. 对每一个子块 R_i, 改变 $\Pi(K,L)$ 和 $L_P(K,L)$, 找到一个最优匹配映射父块 D_j, 使 d_p 达到最小, 则记下 $C_{K,L}$ 和 $h_{K,L}$, $\Pi(K,L)$, $P(K,L)$, 就完成了一个子块 R_i 的编码.

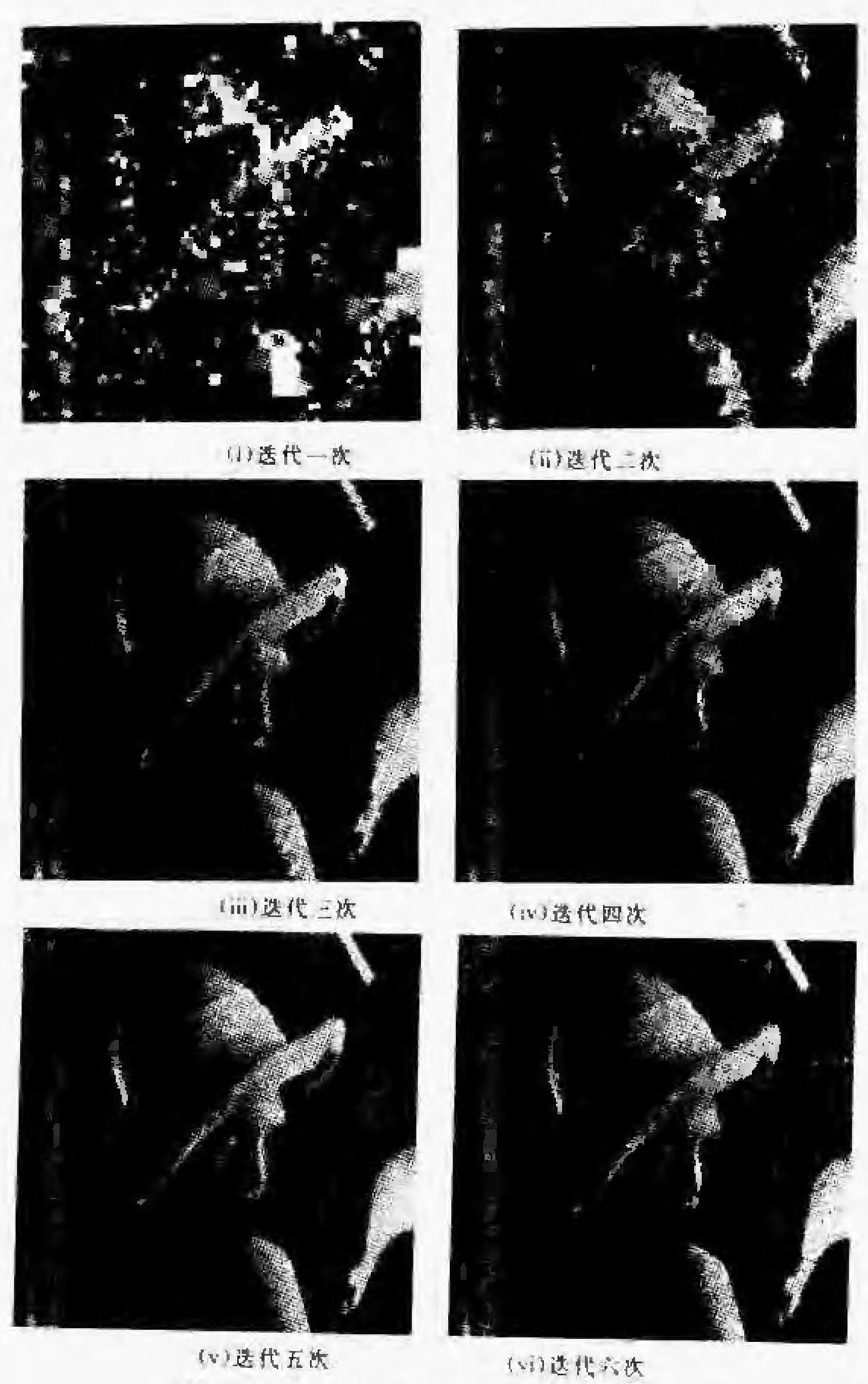
5. 对所有子块, 都分别寻找其对应的父块, 使图象 A 上的每一子块 R_i, 都用其父块来覆盖, 就完成了整幅图象的编码.

图象的解码, 是通过(6,7)式的迭代:

$$\begin{aligned}A&=\sum_{K,L\in R_r}(B^R_{K,L})^*B^R_{K,L}A\\&\approx\sum_{K,L\in R_r}C_{K,L}(B^R_{K,L})^*L_{P(K,L)}A_vB^D_{\Pi(K,L)}A\\&\quad+\sum_{K,L\in R_r}h_{K,L}(B^R_{K,L})^*B^R_{K,L}I).\end{aligned} \tag{6.12}$$

式中 R_r 表示图象中所有子块集.

图 6-1 的(i)～(vi), 表示对一个 256×256 的 Lenna 头像进行解码迭代的过程. 经过 10 次迭代, 完全可达到不动点, 即可看作是原始图象的近似图象.



(i)迭代一次　(ii)迭代二次

(iii)迭代三次　(iv)迭代四次

(v)迭代五次　(vi)迭代六次

图6-1　256×256的Lenna头像分形解码结果

§7 改进的分形图象压缩方法
——四叉树法及其他

采用自动搜索分块的分形编码方法，虽然解决了巴恩斯利子图分割问题，使编码能全自动地进行，但是，分块的搜索需要的计算时间相当可观.

一幅 $2^N\times2^N$ 象素点的带灰阶图象 A，被划成尺度的 $2^R\times2^R$ 象素的子块 R_i，$R_i\cap R_j=\varnothing$，则 A 包括的子块数为

$$n=(2^N\times2^N)/(2^R\times2^R)=(2^{N-R})^2. \tag{7.1}$$

对于每个 R_i，可能搜索的最大自相似块(父块) D_j 的数为 $(2^N-2^D+1)\times(2^N-2^D+1)$，父块的尺寸为 $2^D\times2^D$，$D>R$. 在本书第63页中我们采用了“旋转-反射”的变换有 8 种，这就对每一个 R_i 必须有 $8\times(2^N-2^D+1)\times(2^N-2^D+1)$ 次的匹配计算，通过公式(6.9)、(6.10)、(6.11)，才能找到一个最优的编码参数 $C_{K,L}$，$h_{K,L}$. 其搜索的量级为 $o((2^N)^3)\sim o((2^N)^4)$. 为了降低分形编码时搜索的复杂度，90年代起，很多文章提出了一些改进方法，如固定网络搜索法、分类搜索法、四叉树搜索法和最优搜索法等[24]~[27]，这些文章都为分形编码的实用化开创了途径.

§7.1 自相似块选取上的改进

1. 固定网络法

假设图象 A 有 $2^N\times2^N$ 个象素点. 对每个子块 R_i，需要搜索的可能父块 D_j 数为

$$K=8\times(2^N-2^D+1)\times(2^N-2^D+1). \tag{7.2}$$

将这些可能的父块组成一个父块库，每个子块 R_i 都需要在此父块库中搜索其匹配的自相似父块. 固定网络法是减小 K 的数目，将父块的位置固定在图象的一些网格上，在搜索时不需要对每个可能的父块进行搜索，而只要对那些网格上的父块进行搜索即可. 这

样做的目的，是减小搜索的复杂性，而并不对质量产生多大的影响.

网格有如下几种取法：

i. 网格间隔为 l，均匀分布在图象 A 上，$l>1$，父块的左上角位于网格上. 此时父块数为

$$K_l = \frac{1}{l^2}K. \tag{7.3}$$

如 $l=2$，则父块库中的父块数为原来的 $\frac{1}{4}$.

ii. 网格的间隔为父块的尺寸除以 l. 总的父块数为

$$K_l = \frac{l^2}{2^D \times 2^D} K. \tag{7.4}$$

父块库中的父块数与父块的大小成反比.

iii. 交叉的选取法. 假设在子块与父块搜索自相似的过程中，父块的尺寸是可以变化的，最大的父块尺寸为 $2^{D_{\max}} \times 2^{D_{\max}}$，最小的父块尺寸为 $2^{D_{\min}} \times 2^{D_{\min}}$，父块的左上角位于网格点上，最大的父块网格间隔 $l'_{\max}$ 为最小的父块边长除以 l，最小父块网格间隔 $l'_{\max}$ 为最大的父块边长除以 l：

$$l'_{\max} = \frac{l}{2^{D_{\min}}}, \quad l'_{\min} = \frac{l}{2^{D_{\max}}}. \tag{7.5}$$

2. 邻近搜索法

大量的统计实验表明，对于给定位置的子块 R_i，搜索到其最优匹配自相似块的位置的概率分布几乎都集中在子块 R_i 附近. 为了减少搜索的父块数目，在编码时给定一个允许误差 ε，这里 ε 是一个指定的小正数，对每个给定子块，首先搜索的是其邻近区域的父块，(6.7)式中的近似式改为

$$B^R_{K,L}A \approx C_{K,L}L_{P(K,L)}A_v B_{\Pi'(K,L)}^{D}A + h_{K,L}B^R_{K,L}I. \tag{7.6}$$

$\Pi(K,L)$ 改为 $\Pi'(K,L)$，$\Pi'(K,L) \Diamond (K+\Delta K, L+\Delta L)$，$\Delta K$、$\Delta L \in N_r$，$N_r$ 为坐标 (K,L) 的邻近，其象素点的距离为 r. $r=1$ 是周围 8 个象素坐标位置，$r=2$ 是周围 24 个象素坐标位置.

搜索时，逐步扩大 r 的值，直到 $d_p<\varepsilon$ 为止. 这样可避免对父块库中的全体父块进行搜索，同时由于自相似块比较集中在子块 R_i 的周围，只要搜索少量的父块就能找到要求误差的自相似父块.

3. **分类的父块库搜索法**

根据父块的特征，预先把父块分为几类. 然后，对不同的类的子块，分别在它所对应的父块类中搜索. 类分得越精细，其父块搜索的范围就越小. 这就可以节约搜索的时间，得到有效的分形编码. 分类有两种：

i. 利用父块的几何特征进行分类[21]

假设父块库 $\mathbb{D}$ 是由 K 个父块组成，每个父块具有 $m=2^D\times 2^D$ 个象素点. 对给定一个父块 D_j，$j=1,2,\cdots,K$，它的每个象素点上的灰阶值为 b_i^j，其中 $i=1,2,\cdots,2^D\times 2^D$. 根据灰阶的分布的几何特征，把 K 个父块分为平坦块 D^s、边缘块 D^e、中间纹理块 D^m 这三类. 父块库 $\mathbb{D}$ 满足：

$$\mathbb{D}=D^s\cup D^e\cup D^m,$$

且

$$D^s\cap D^e=\varnothing,\ D^e\cap D^m=\varnothing,\ D^m\cap D^s=\varnothing. \tag{7.7}$$

D^s、D^e 和 D^m 分类的依据，是每个父块灰阶的均值和方差. 一个父块 D_j 它的灰阶均值为

$$\mu_j=\frac{1}{m}\sum_{i=1}^{m}b_i^j, \tag{7.8}$$

其方差为

$$\mathrm{var}_j=\frac{1}{m}\sum_{i=1}^{m}(b_i^j-\mu_j)^2. \tag{7.9}$$

给定阈值 T_1、T_2、T_3，满足 $T_1\ll T_3$，$T_2<T_3$. 对于父块 D_j 的分类是：

当 $\mathrm{var}_j<T_1$，且 $\sum\limits_{\substack{i_1\neq i_2\\ i_1,i_2\in D_j}}|b_{i_1}^j-b_{i_2}^j|<T_2$ 时，为 D^s 类——平坦类. 此

类父块中，每个象素之间的灰度差很小，几乎可看作近似均值 μ_j. 如图片中的蓝天，平静的大海，即属于同一目标上的块.

当 $\text{var}_j > T_3$，且 $\sum_{\substack{i_1 \neq i_2 \\ i_1, i_2 \in D_j}} |b^j_{i_1} - b^j_{i_2}| > T_2$ 时，为 D' 类——边缘类. 此类父块中，包含了物体的边缘，有明显的黑白反差.

当 $T_1 < \text{var}_j < T_3$，且 $\sum_{\substack{i_1 i_2 \in D_j \\ i_1 \neq i_2}} |b^j_{i_1} - b^j_{i_2}| > T_2$ 时，为 D'' 类. 此类多为纹理性的块，如树皮，衣服的花纹等，它的方差不像边缘块那么大，但象素点之间的灰度差距仍比较大.

阈值 T_1、T_2、T_3 是由实验决定的. 一旦阈值定了，每个父块库中的父块都可计算出属于它的类. 给定一个子块 R_i，同样可按照上面的标准计算出这个子块是属于那一类的，然后根据不同类的子块，在其对应的父块类中搜索其相应的自相似块. 这样不仅减少了搜索的父块数目，而且对于平坦块的处理也比较简单.

如子块 R_i 为平坦类的，就不需要搜索其对应的父块，可用子块 R_i 本身的平均灰阶值来近似. 即

$$B^R_{K,L}A \approx h_{K,L} B^R_{K,L} I, \tag{7.10}$$

式中 $h_{K,L} = \mu_{B^R_{K,L}A}$，$C_{K,L} = 0$，$(K,L)$ 为 R_i 在图 A 中左上角的坐标.

对于子块 R_i 属于边缘类或纹理类的情况，只要分别在父块库中的 D' 和 D'' 类中搜索，同时按(6.9)、(6.10)、(6.11)式计算，找到其自相似的匹配块. 在三种不同类均等的条件下，搜索的范围只是原来的 1/3，甚至更少，因为平坦块就不用搜索了.

ii. 四分之一灰度排列的分类法[22]

设 D_j 为父块库中任一父块，$j=1,2,\cdots,K$，具有 $2^D \times 2^D$ 的象素点，把它分割为上左、上右、下左、下右四个部分，并依次记为 D_{j1}、D_{j2}、D_{j3}、D_{j4}，其每 $\frac{1}{4}$ 块的灰阶值为 $b_1^{j_1} b_2^{j_1}, \cdots, b^{j_1}_{n/4}$；$b_1^{j_2} b_2^{j_2}, \cdots, b^{j_2}_{n/4}$；$b_1^{j_3} b_2^{j_3}, \cdots, b^{j_3}_{n/4}$；$b_1^{j_4} b_2^{j_4}, \cdots, b^{j_4}_{n/4}$. 每个 $\frac{1}{4}$ 块的均值 μ_{j_i} 和方差 var_{j_i} 分别为：

$$\mu_{j_1}=\sum_{i=1}^{m/4}b_i^{j_1},\ \mu_{j_2}=\sum_{i=1}^{m/4}b_i^{j_2},\ \mu_{j_3}=\sum_{i=1}^{m/4}b_i^{j_3},\ \mu_{j_4}=\sum_{i=1}^{m/4}b_i^{j_4};\tag{7.11}$$

$$\mathrm{var}_{j_1}=\frac{4}{m}\sum_{i=1}^{m/4}b_i^{j_1}-\frac{4}{m}\mu_{j_1},\ \mathrm{var}_{j_2}=\frac{4}{m}\sum_{i=1}^{m/4}b_i^{j_2}-\frac{4}{m}\mu_{j_2},$$

$$\mathrm{var}_{j_3}=\frac{4}{m}\sum_{i=1}^{m/4}b_i^{j_3}-\frac{4}{m}\mu_{j_3},\ \mathrm{var}_{j_4}=\frac{4}{m}\sum_{i=1}^{m/4}b_i^{j_4}-\frac{4}{m}\mu_{j_4}.\tag{7.12}$$

对于任一个父块 D_j，都可通过"反射-旋转"变换，使它的$\frac{1}{4}$块的均值排列属于下面三种情况之一，从而父块也被分为三大主类：

主类 1： $\mu_{j_1}\geqslant\mu_{j_2}\geqslant\mu_{j_3}\geqslant\mu_{j_4}$，

主类 2： $\mu_{j_1}\geqslant\mu_{j_2}\geqslant\mu_{j_4}\geqslant\mu_{j_3}$，

主类 3： $\mu_{j_1}\geqslant\mu_{j_4}\geqslant\mu_{j_2}\geqslant\mu_{j_3}$.

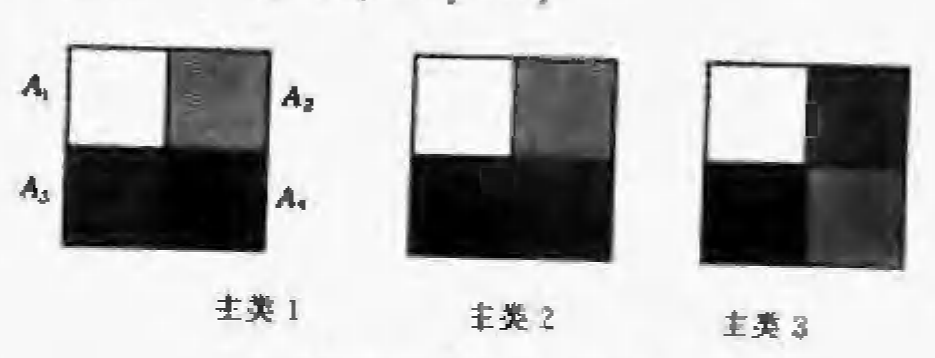


图 7-1 按$\frac{1}{4}$块的均值分类情况

对每个主类，又按$\frac{1}{4}$块的方差 var 的大小分为24个副类，这就把父块库分为 72 类. 当然，方差 var 的副类数目可根据方差的等级，可分多或分少，这完全由问题而定.

在分形编码时，对每一个子块 R_i，先按 D_j 的形式那样分为四个$\frac{1}{4}$块，按照(7.11)计算均值，通过 $L_{P(K,L)}$算子寻找对应的主类的父块，然后，按式(7.12)计算$\frac{1}{4}$块的方差，找到对应父块的副类；最后，用(6.9)、(6.10)、(6.11)式找到其自相似块. 由於搜索是在$\frac{1}{72}$

的父块库上进行，大大缩短了搜索时间和计算复杂度.

§7.2 自适应的四叉树编码方法

上述方法讨论的父块和子块，都是在固定尺寸上进行的. 其实，父块和子块的尺寸，对于分形图象编码的质量和压缩比的影响很大. 对父块为 $2^D\times 2^D$ 象素，子块为 $2^R\times 2^R$ 象素，如 D 和 R 很大，虽然父块库中的父块数目比较小，但很难找到子块与父块之间的相似关系，即使利用(6.9)～(6.11)式求出最优匹配系数，但原图 A 与拼贴一次仿射变换后的图的误差很大，按§2中的拼贴定理，将来解码的图象的质量也很差. 但如 R 和 D 尺寸很小，最小的情况一个子块为一个象素 $2^0\times 2^0$，父块为 4 个象素，那么总可以找到非常好的自相似匹配系数，此时图象编码的仿射变换数目巨增，并不能达到图象压缩的目的.

弗歇尔和简库恩[20],[24],[26]等人提出了根据允许的误差 ϵ 而改变子块和父块的尺寸大小的做法，既能对于大尺寸的子块和父块之间的相似性做合理的选取仿射变换系数，而对大尺寸达不到要求的子块，改变其边长，从 2^R 到 $2^{R-1}, 2^{R-2}, \cdots$，相应的父块也可同时缩小，父块与子块之间的匹配尺寸是以要求误差 ϵ 和压缩比来进行的，这种方法称为四叉树的自适应方法.

设 ϵ 为一个分形压缩编码过程中，父块与子块之间匹配的允许误差，$2^{R_{max}}$、$2^{R_{min}}$分别为最大子块边长和最小子块边长，能达到压缩的最小于块尺寸为 $2^{R_{min}}\times 2^{R_{min}}$，其四叉树方法的步骤为：

i. 原始图象 A，分割成不相交叠的 $2^{R_{max}}\times 2^{R_{max}}$ 的子块 R_i 作为初始的分割，R_i 的左上角位于(K,L)的位置上，用取块符表示 $B^{R_{max}}_{K,L}A$.

ii. 搜索其相应的父块，位置在 $\pi(K,L)$上，采用"反射-旋转"操作，和平均抽样操作搜索的相应父块为 $L_{P(K,L)}A_vB^{D_{max}}_{\pi(K,L)}A$. 设子块与变换后的父块的灰阶分别为 $a_1, a_2, \cdots, a_m, b_1, b_2, \cdots, b_m$.

子块与变换后的父块之间的误差为：

$$d_p(B_{K,L}^{R\max}A,(C_{K,L}^{R\max}L_{p(K,L)}A_v^{D_{\max}-R_{\max}}B_{\pi(K,L)}^{D\max}A+h_{K,L}^{R\max}I))$$

$$=\frac{1}{m}\sum_{i=1}^{m}(C_{K,L}^{R\max}b_i+h_{KL}^{R\max}-a_i^{R_{\max}})^2. \qquad (7.13)$$

令$\frac{\partial d_p}{\partial C_{KL}}=0$、$\frac{\partial d_p}{\partial h_{KL}}=0$，求出$C_{K,L}^{R\max}$，$h_{K,L}^{R\max}$. 按式(6.11)式计算出$d_{p_{\min}}$.

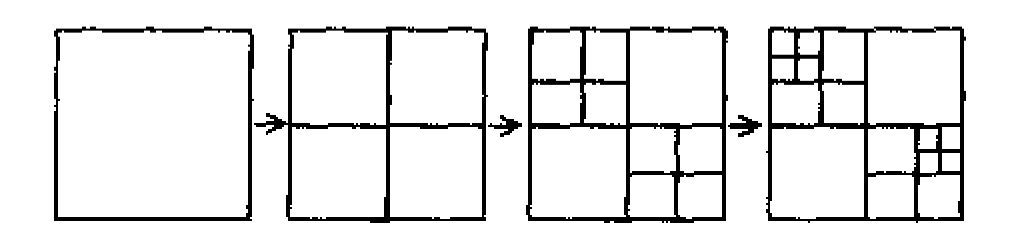
iii. 如果$d_{p_{\min}}<\varepsilon$，则记下(K,L)，$\pi(K,L)$ $P(K,L)$，$C_{K,L}^{R\max}$，$h_{K,L}^{R\max}$，就完成了对R_i的编码. 否则，转到iv.

iv. 如果$d_{p_{\min}}>\varepsilon$，则把子块R_i再划分为四块，父块也相应划为4块. 此时取$t=t+1$，子块取块仍用$B_{K,L}^{R\max-t}A$表示. 对父块的变换相应地为$L_{P(K,L)}A_vB_{\pi(K,L)}^{D\max-t}A$，继续采用(7.13)式，此时，

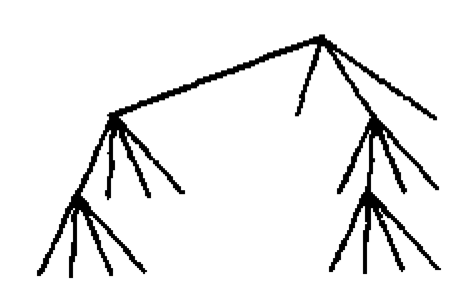
$$d_p=(B_{K,L}^{R\max-t}A,(C_{K,L}^{R\max-t}L_{P(K,L)}A_v^{D_{\max}-R_{\max}}B_{\pi(K,L)}^{R\max-t}A+h_{K,L}^{R\max-t}I))$$

$$=\frac{4}{m}\sum_{i=1}^{m/4}(C_{K,L}^{R\max-t}b'_i+h_{K,L}^{R\max-t}-a'^{R_{\max}-t}_i)^2. \qquad (7.14)$$

这里的b_i'、a_i'都是在四叉树划小时的块中的灰度值，并再按(6.9)～(6.11)求出$d_{p_{\min}}$.



(i) 四叉树分割过程



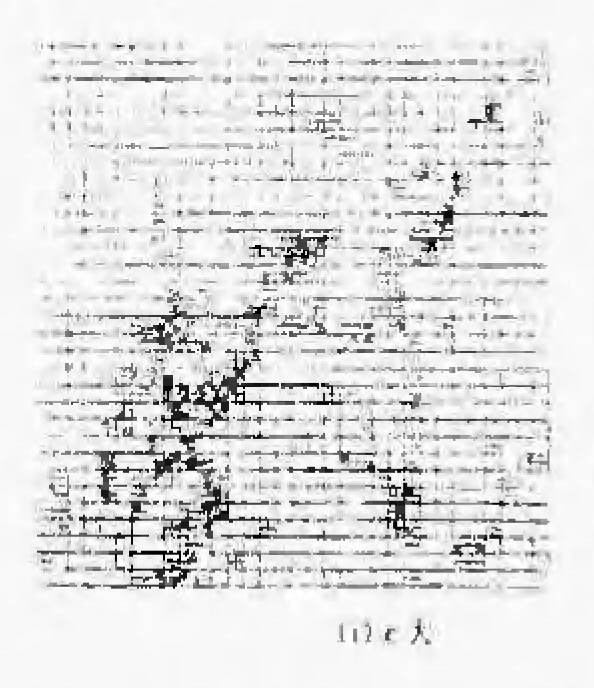
(ii) 对应(i)的四叉树数

图 7-2 四叉树分割的过程

v. 重复iii和iv，直到所有的大小子块与相适应的父块之间自相似变换误差都小于ε为止. 图7-2表示一个四叉树的分割过程，分割时，只有对$d_{p_{\min}}>\varepsilon$的那些子块才进行四叉树的划分. 因此，

在图象变化复杂的地方，块划得比较小；而平坦的地方，块比较大.

图7-3表示四叉树分割法对 Lenna 头像的编码和解码的结果.图7-3(i)、(ii)分别为 ε 比较大的四叉树分割和 ε 小的四叉树分割图，其解码结果为图7-3(iii)、(iv). 当 ε 大时图7-3(iii)其压缩比高，但它的图象质量无论从客观标准 PSNR 还是主观标准看，都比图7-3(iv)差.



(i) ε 大　　(ii) ε 小



(iii) 图i解码后的结果　　(iv) 图ii解码后的结果

压缩比 37.2　　压缩比 17.8

PSNR = 28.3 db　　PSNR = 30 db

图7-3　不同 ε 时 Lenna 头像四叉树分割的情况

§7.3 其他的分形编码改进方法

1. 在父块与子块匹配计算量上的改进

i. 函数方法匹配

[25]提出一个思想：不直接比较子块与父块之间的自相似映射关系，而是把子块与父块都与一个单位矢量进行比较，只对与单位矢量关系接近的那些父块才进行搜索. 为了说明此方法，把从图象 A 中取出的子块 $B^{R}_{K,L}A$ 和通过取块、变换后的父块($L_{P(K,L)}A, B^{D}_{\kappa(K,L)}A$)排列为一维的矢量，记为 $\boldsymbol{V}_R$ 和 $\boldsymbol{V}_D$，这里 $\boldsymbol{V}_R$、$\boldsymbol{V}_D \in \mathbb{R}^m$. 记 $\boldsymbol{U}$ 为一个单位矢量，$\boldsymbol{U} \in \mathbb{R}^m$. 设 $\eta(\boldsymbol{V})$ 为对 $\boldsymbol{V}$ 取均值归一化的矢量.

$$\eta(\boldsymbol{V}) = \frac{1}{\sqrt{\mathrm{var}(\boldsymbol{V})}}(\boldsymbol{V}_1 - \overline{\boldsymbol{V}}, \boldsymbol{V}_2 - \overline{\boldsymbol{V}}, \cdots, \boldsymbol{V}_m - \overline{\boldsymbol{V}}),\ \overline{\boldsymbol{V}} \in \mathbb{R}^m$$

$$\overline{\boldsymbol{V}} = \frac{1}{m}(\boldsymbol{V}_1 + \boldsymbol{V}_2 + \cdots + \boldsymbol{V}_m),$$

$$\mathrm{var}(\boldsymbol{V}) = \sum_{k=1}^{m}(\boldsymbol{V}_k - \overline{\boldsymbol{V}})^2.$$

定理 1[29] 让 δ 为选择的允许误差，$\delta>0$；$\boldsymbol{V}_R$、$\boldsymbol{V}_D$ 为子块与变换后的父块组成的矢量，$\boldsymbol{U}$ 为单位矢量 $\boldsymbol{V}_R$、$\boldsymbol{V}_D$ 和 $\boldsymbol{U} \in \mathbb{R}^m$，且有 $\langle \boldsymbol{V}_R, \eta(\boldsymbol{V}_R)\rangle > \delta$. 假定(6.11)式中的 $d_{p_{\min}} < \delta$，则有

$$|\langle \eta(\boldsymbol{V}_R), \boldsymbol{U}\rangle - \langle \eta(\boldsymbol{V}_D), \boldsymbol{U}\rangle| \leqslant \sqrt{2 - \sqrt{1 - \frac{\delta^2}{\langle \boldsymbol{V}_R, \eta(\boldsymbol{V}_R)\rangle^2}}}. \tag{7.15}$$

上式中，$\langle \boldsymbol{V}_R, \eta(\boldsymbol{V}_R)\rangle$ 为内积表达式. 定理 1 的证明可见[31]. 当内积满足 $\langle \boldsymbol{V}_R, \eta(\boldsymbol{V}_R)\rangle \geqslant \delta$ 时，意味着 $\mathrm{var}(\boldsymbol{V}_R) \geqslant \delta^2$，具体函数算法为：

①选择一个允许误差 δ，和一个单位矢量 $\boldsymbol{U}$；

②对每个父块，进行预处理 $\langle \eta(\boldsymbol{V}_D), \boldsymbol{U}\rangle$；

③对每个子块，计算 $\langle \boldsymbol{V}_R, \eta(\boldsymbol{V}_R)\rangle$ 和(7.15)式右面的值，如果 $\langle \boldsymbol{V}_R, \eta(\boldsymbol{V}_R)\rangle < \delta$，则不需要在父块库中进行搜索；

④如果$\langle V_R,\eta(V_R)\rangle\geqslant\delta$，计算$\langle\eta(V_R),U\rangle$. 对于不满足不等式(7.15)的父块，不予计算.

如有可能，可考虑多个单位矢量U. 这种方法也是不需要在父块库中全面的匹配搜索，而只要对满足(7.15)的不等式的部分父块进行匹配搜索. 这个方法也可以推广到更一般化的形式. 设存在一个函数$F:\mathbb{R}^m\to\mathbb{R}$，如果$|F(V_D)-F(V_R)|\leqslant\varepsilon_R$，则意味着子块矢量$V_R$能够与$V_D$匹配得比较好. 如果父块库是由$\{V_{D1},V_{D2},\cdots,V_{DN}\}$组成，它的匹配搜索步骤为

①对每一个$\{V_{D1},V_{D2},\cdots,V_{DN}\}$，计算$F(V_D)$；

②把$F(V_{Di})$进行排队，使

$$F(V_{D1})\leqslant F(V_{D2})\leqslant\cdots\leqslant F(V_{DN});$$

③对每子块矢量V_R，计算$F(V_R)$和公式(7.15)的误差上限，找到一个排序数k_0,k_1，使$|F(V_R)-F(V_{Dk})|\leqslant\varepsilon_k$当$k_0\leq k\leq k_1$时成立；

④对于所有的在$k_0\leqslant k\leqslant k$中的父块矢量$V_{Dk}$，进行(6.9)、(6.10)、(6.11)的计算，并按d_p达到最小的父块进行系数的编码. 这就大大简化了计算量.

ii. 多分辨率的方法

[28]、[30]、[31]、[32]提出一种进行搜索父块与子块相匹配的塔型的多分辨率方法. 把原始图象A分为几种分辨率，设A具有$2^N\times2^N$个象素点，为最高分辨率. $A_{\frac{1}{4}}$、$A_{\frac{1}{16}}$、$A_{\frac{1}{34}}$，…，$A_{\frac{1}{2^P\times2^P}}$分别为具有不同分辨率的图象，它们的象素点为$2^{N-1}\times2^{N-1}$，$2^{N-2}\times2^{N-2}$，$2^{N-3}\times2^{N-3}$，…，$2^{N-P}\times2^{N-P}$.

$$A_{\frac{1}{2^k\times2^k}}(i,j)=\frac{1}{4}\sum_{l,s=0}^{1}A_{\frac{1}{2^{k-1}\times2^{k-1}}}(2i+l,\ 2j+s).$$

$$k=1,2,\cdots,P,\ \ 0\leqslant i,j<2^k$$

定理 2 设$R^{(k)}$和$D^{(k)'}$表示在分辨率k时子块和已变换后的父块，则匹配的误差$d_p(D^{(k+1)'};R^{(k+1)})\geqslant d_p(D^{(k)'};R^{(k)})$.

对父块的搜索先在分辨率低的情况下进行，当找到其匹配块

的情况下，再从分辨率高的图象上去找其相适应的仿射变换．定理2表明，如果一个子块在低分辨率时不能找到其匹配的父块，则在高分辨率时就不可能找到其相应的匹配块．四叉树的方法先在低分辨率时进行，然后从低分辨率找到的那些匹配块中，再提高分辨率进行搜索其匹配父块．多分辨率方法使搜索的目标范围缩小，也简化了搜索时的计算量．

2．其他的改进方法

i．划块形状的改变

以上讨论的子块、父块的划分，是建立在正方形的基础上的．但是，正方形并不是最好的分块形式．有些情况下，矩形、三角形、菱形都可作为图象划分的形式．虽然这些划分可能加重了计算量，但有时为了提高图象解码后的质量，可采用这些非正方形的块的分割．下面以矩形划分为例．[33]

图象 A 首先被划分为矩形．如给定一个矩形子块 R_i，找不到其相适应的父块 D_i，则四叉树的方法被改为用两种分割进行：一个为平行分割为两个矩形，另一个为垂直分割为两个矩形，然后再进行匹配查找．其分割的标准是建立在图象的灰阶的均匀性上．

对于一个子块 R_i，它的灰阶函数为 $f_{R_i}(i_x,i_y)$，定义在 $0\leqslant i_x<m$，$0\leqslant i_y<n$，$m\neq n$ 的区域上，可计算其垂直方向和水平方向上的不均匀性情况：

$$h_{i_y}=\frac{\min(i_y,\ n-i_y-1)}{n-1}\Big(\sum_{i_x}f_{R_i}(i_x,i_y)-\sum_{i_x}f_{R_i}(i_x,i_y+1)\Big),$$

$$v_{i_x}=\frac{\min(i_x,\ m-i_x-1)}{m-1}\Big(\sum_{i_y}f_{R_i}(i_x,i_y)-\sum_{i_y}f_{R_i}(i_x+1,\ i_y)\Big).\qquad(7.16)$$

这两个不均匀数量的最大值 $\max(h_{i_y},v_{i_x})$，决定分割的方向和位置，并以这种矩形分割的子块寻找有关父块．这种方法对于分割的

计算不像四叉树那么简单，却能带来比较好的解码图象的质量（如图7-4）. 这种分割方法称为平行和垂直划分（HV 划分）.

划块的形状还有三角形或菱形，这里不细介绍，可参见[31]、[35]. 更复杂形状的分割，如 F. Thomas 和 F. Deravi 提出的一种方法：首先把图象分割为 4×4 或 8×8 的方块，而相邻的块如果均值和偏差相似，通过合并，建立起大的块，因而最终可能形成一些更大的块，这些大块的自相似变换的参数计算出来就可恢复图象. 但是在图象的形状和结构不规则和比较复杂时，这种方法不太容易找到匹配块，于是在操作时采用了几个层次，把小的块称为"原子块"，先用简单的搜索方法找到"原子块"的最优匹配块；然后搜索这个匹配块周围的块，看是否可以合并到此块中来，这种扩张一直到其失真判据不满足，然后对另一个"原子块"，继续进行，直到整幅图象都编码完毕. 由于"原子块"的尺寸比较小，而合并进来的邻子块和父块的形状可以从各个方向上进行拼贴，这样对图象的拼贴就更完美. 如图7-5所示，这种方法的详细介绍可参见[35].



图7-4　HV 分割的结果

ii. 利用进化算法进行子块对父块的搜索[36]

进化算法是一种随机搜索方法，它并不需要像§6 那样进行穷举的搜索，找到对应于一个子块相匹配的父块，而是通过随机抽取少量的父块集（称为进化集），在进化集中进行匹配搜索，找到的匹配较优的父块，进行传种、扩大；而对匹配较差的父块进行淘汰. 通过几次迭代，那些优的父块能逐步进化，达到较快的搜索到最优的匹配父块的目的. 具体算法如下：

① 对给定一个正方形子块 R，随机取出 q 个父块，用标准分形编码方法找到与 R 匹配的最好的 q_s 个父块，$q_s \geqslant 1$，每个父块对

应的尺度参数 $C_{K,L}^{R}$ 和偏移参数 $h_{K,L}^{R}$ 都记录下来(作为参数).



图7-5 进化与拼合方法结合的结果

② q_d 个父块所有的参数形成进化集(population),它们由匹配得较好的一些父块组成.

③ 进化算法是在进化集中的父块,通过随机的合并其周围两个相邻的父块,或直接计算周围两个随机的相邻父块,此时 q_d 的父块变成 q_d+2q_d 个父块的参数集.

④ 通过对 $3q_d$ 个父块对给定的子块 R 之间的符合度函数的测试(这里的符合度函数就是(6.11)中的 d_{min},并进行排队,找出最优的 q_d 个父块,放到进化集内.

⑤ 继续进行 ③、④ 步多次,就可找到最优匹配块.

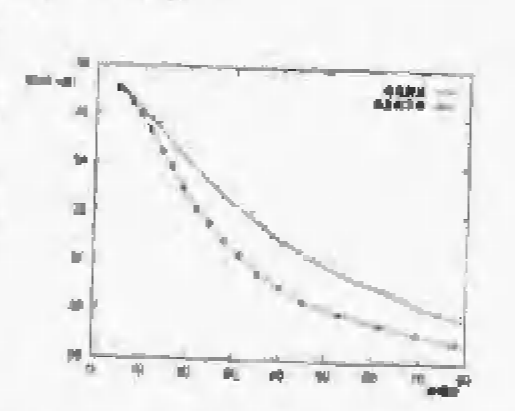


图7-6 进化算法与四叉树方法对比

这种方法与标准的搜索方法相比,它的搜索速度和质量有一定的提高.图7-6所示,是用进化方法和合并方法组合的结果曲线.在同一个压缩比下,其图象质量有所提高;同时,由于进化算法不需要穷举搜索,其搜索速度也大大加快.

§8 运动图象的分形压缩方法

早在1989年,巴恩斯利[38]从理论上提出了IFS可推广到活动图象.只是它的理论是建立在分形空间的点集上,讨论的是正问题,即从已知的仿射变换中迭代出自然图象.如对于这些仿射变换系数略加控制,此自然图象可以活动起来,好像是微风吹动树后的

叶子飘动．如巴恩斯利没有发表静态图象压缩方法一样，他也没有发表从活动的自然图象中自动抽取仿射变换系数的方法，因而基于块的仿射变换系数计算方法，仍为运动图象压缩的主要途径．

1．**活动图象的自仿射映射**

§2的定理2告诉我们，一个$\{\mathscr{X},w_1,w_2,\cdots,w_n\}$是$(\mathscr{X},d)$上的IFS，$W:\mathscr{H}(\mathscr{X})\to\mathscr{H}(\mathscr{X})$是度量空间上的压缩变换，$W(B)=\bigcup_{i=1}^{n}w_i(B)$，$\forall B\in\mathscr{H}(\mathscr{X})$．存在唯一的不动点$\overline{A}=W(\overline{A})$，此不动点可选择任意初值通过多次迭代而得到：$\overline{A}=\lim\limits_{k\to\infty}W^k(B)$．此即不动点定理，是分形解码的基础．同时在分形编码中按照§2的定理3的拼贴定理，即给定集合$L\in\mathscr{H}(\mathscr{X})$和$\varepsilon>0$，如找到一个IFS $\{\mathscr{X},w_1,w_2,\cdots,w_n\}$，使其豪斯多夫距离$h(L,\bigcup_{i=1}^{n}w_i(L))\leqslant\varepsilon$，则$h(L,\overline{A})\leqslant\frac{\varepsilon}{1-S}$，$S$为压缩因子，这是图象编码的原理．

但在有些图象中存在一些特殊的点集，如树叶的柄，它是由一些特殊变换集变换得到的，为了讨论在点与点之间的仿射变换中的这些特殊点，巴恩斯利提出了凝聚集的概念．

定义1　设$(\mathscr{X},d)$是一个度量空间，$C\in\mathscr{H}(\mathscr{X})$．定义一个变换$w_0$：$\mathscr{H}(\mathscr{X})\to\mathscr{H}(\mathscr{X})$，$w_0(B)=C$，$\forall B\in\mathscr{H}(\mathscr{X})$．那么，$w_0$称为**凝聚变换**，而$C$称为与$w_0$有关的**凝聚集**(condensation set)．

定义2　设$\{\mathscr{X},w_1,w_2,\cdots,w_n\}$是带有压缩因子$0\leqslant S<1$的IFS，设$w_0$：$\mathscr{H}(\mathscr{X})\to\mathscr{H}(\mathscr{X})$是凝聚变换，那么$\{\mathscr{X},w_0,w_1,w_2,\cdots,w_n\}$称为“带有凝聚和压缩因子$S$的IFS”．

定理1（扩展的不动点定理）　设$\{\mathscr{X},w_i:i=0,1,2,\cdots,n\}$是带有凝聚和压缩因子$S$的变换，那么变换$W:\mathscr{H}(\mathscr{X})\to\mathscr{H}(\mathscr{X})$定义为

$$W(B)=\bigcup_{i=0}^{n}w_i(B),\qquad \forall B\in\mathscr{H}(\mathscr{X})\quad(8.1)$$

是一个完备度量空间$(\mathscr{H}(\mathscr{X}),h(d))$上的带有压缩因子$S$的压缩映射，它的唯一不动点$\overline{A}\in\mathscr{H}(\mathscr{X})$，满足

$$\overline{A}=W(\overline{A})=\bigcup_{i=0}^{n}W_i(\overline{A}), \tag{8.2}$$

可以通过迭代来达到：$\overline{A}=\lim_{k\to\infty}(W^0)^{(k)}(B),\forall\ B\in\mathscr{H}(\mathbb{X})$.

这里的$(W^0)^{(k)}$，表示带有凝聚集的变换，k为迭代次数. 这与§2中的定理2相比，多了一个w_0有关的凝聚集.

定理2（扩展的拼贴定理） 设$(\mathbb{X},d)$是完备的度量空间，$L\in\mathscr{H}(\mathbb{X})$是一个给定分形集，$\varepsilon<0$. 选择一个带有凝聚变换的IFS $\{\mathbb{X};w_0,w_1,\cdots,w_n\}$，具有压缩因子$0\leqslant S<1$. 如果满足：

$$h(L,\ \bigcup_{i=0}^{n}w_i(L))\leqslant\varepsilon, \tag{8.3}$$

那么

$$h(L,A)\leqslant\varepsilon/(1-S), \tag{8.4}$$

上式中的A是IFS的不动点(吸引子).

带有凝聚的仿射变换比一般的变换多了一个w_0的变换. 这个变换使自仿射变换在迭代中多了一个约束条件，使变换容易控制.

图8-1表示一个IFS的吸引子，是用了三维进行变换的$\{\mathbb{R}^3;w_i,i=1,2,3,4\}$，吸引子在区域$\{(x_1,x_2,x_3)\in\mathbb{R}^3;\ -10\leqslant x_1\leqslant 10,0\leqslant x_2\leqslant 10,-10\leqslant x_3\leqslant 10\}$. 用下面四个自仿射变换进行迭代，得到的不动点犹如一棵树被风吹起来的情况，这是一般的二维自仿射变换做不到的.

$$W_1\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0&0&0\\0&0.18&0\\0&0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}+\begin{pmatrix}0\\0\\0\end{pmatrix},$$

$$W_2\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0.85&0&0\\0&0.85&0.1\\0&-0.1&0.85\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}+\begin{pmatrix}0\\1.6\\0\end{pmatrix},$$

$$W_3\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}0.2&0.2&0\\0.2&0.2&0\\0&0&0.8\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}+\begin{pmatrix}0\\0.8\\0\end{pmatrix},$$

$$W_4\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=\begin{pmatrix}-0.2 & 0.2 & 0\\0.2 & 0.3 & 0\\0 & 0 & 0.8\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}+\begin{pmatrix}0\\0.8\\0\end{pmatrix}.$$

定理3 设 $\{X,d\}$ 是一个度量空间，$\{X;w_0,w_1,\cdots,w_n\}$ 是带凝聚压缩因子 S 的 IFS. 让 W_i 连续地与一个参数 p 有关，$p\in P$，P 是一个完备度量空间. 那么，吸引子 $A(P)\in H(X)$ 也连续与 p 相关.

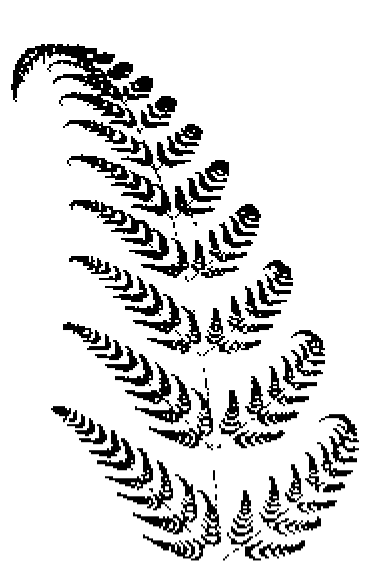


图 8-1 被风吹起的叶子

定理3告诉我们，只要变化参数 p，可以使吸引子 A 发生变化，可以通过参数 p 来控制 IFS 的吸引子. 如果 p 的变化是连续的，则吸引子 $A(p)$ 也连续变化. 这样可以使静止的图象在 p 变化时发生了活动的作用. 图8-2中的六幅图象，表示在参数 p 变化后得到的不同的吸引子，如树被风刮起的情况.

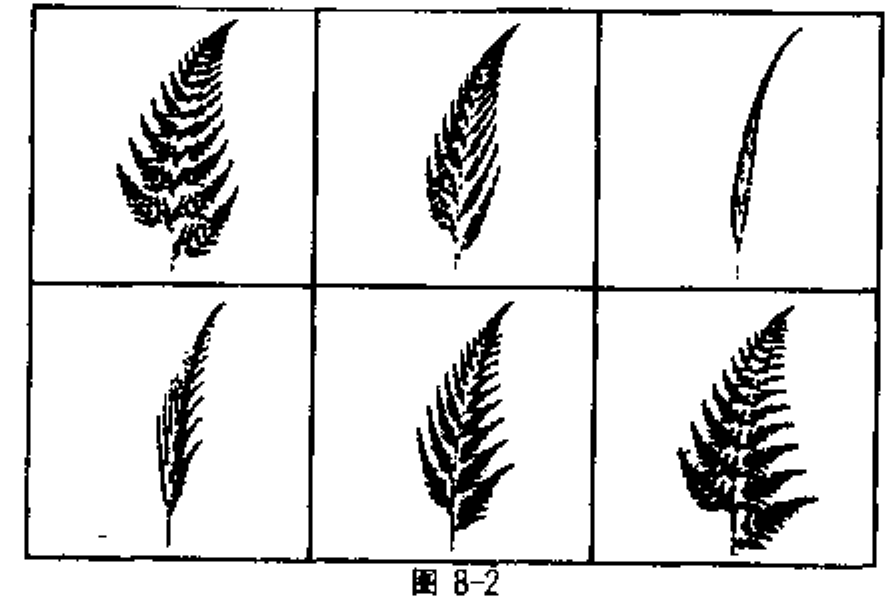


图 8-2

2. 有灰阶的运动图象压缩编码方法

巴恩斯利给出的是分形点集在参数变化时的吸引子变化情况，是正变换：即已知自仿射变换的参数和控制参数 p 的变化，而得到的是二维图象活动的情况. 但求其逆问题，即已知活动图象而

如何找到其变换 W 系数,这就是一个难题.但定理1、定理2和定理3给活动图象的分形编码提供了理论依据.

i. 完全基于静态图象的活动图象分形编码

运动图象可理解为一系列静止图象按照一定次序播放形成的图象序列.其中,每一幅"静止"图象称为**帧**(frame),对静止图象编码称为**帧内编码**,利用帧与帧之间的相关性进行编码称为**帧间编码**.对于彩色图象,它由 V_R、V_B、V_G 三种基色组成(红、蓝、绿).在实际工作中,把三种基色转换为亮度 V_Y 和色差 V_{R-Y}、V_{B-Y},变换公式为

$$
\begin{aligned}
V_Y &= 0.3V_R + 0.59V_G + 0.11V_B, \\
V_{R-Y} &= 0.70V_R - 0.59V_G - 0.11V_B, \\
V_{B-Y} &= -0.30V_R - 0.59V_G + 0.89V_B.
\end{aligned}
\tag{8.5}
$$

信号的能量主要集中在 V_Y 上,V_{R-Y} 和 V_{B-Y} 能量很小,从而可得到更大的压缩.它与静止图象压缩不同的主要有下面几点:

1. 运动检测:将图象分割为静止区和运动区两部分;
2. 运动估计:对运动区进行位移估计;
3. 运动预测:用位移估计进行运动区的预测编码;
4. 误差补偿:对预测误差进一步编码.

完全基于静态分形编码的运动图象编码,是在一个给定帧内采用传统的分块的分形编码,而帧间编码中采用运动检测和子块与父块匹配的方式,只是父块不是在本帧中找.假设 $R(i,j,t-1)$ 表示前一帧的一个子块,通过帧内的分形编码找到在自己帧内的匹配父块 $D(k,l,(t-1))$,这里的 i,j,k,l 表示子块和父块的位置,其变换的系数也被存储下来.设 $R(i,j,t)$ 是当前帧在相同空间位置上的子块.如果令 $u(i,j,t)$ 和 $u(i,j,t-1)$ 是在 t 和 $(t-1)$ 时刻子块 $R(i,j)$ 在 i,j 位置上的灰阶函数.为了简单起见,我们只考虑亮度分量 V_y,至于色度分量 $V_{R-Y}V_{B-Y}$,也可用类似的方法处理.

令 $T_t>0$ 为预定的阈值,MSE 为两帧的子块之间的误差

$$\begin{aligned}\text{MSE} &= d(R(i,j,t),R(i,j,t-1)) \\ &= \frac{1}{2^K \times 2^K}\sum_{i=1}^{2^K}\sum_{j=1}^{2^K}(u(i,j,t)-u(i,j,t-1))^2.\end{aligned} \tag{8.6}$$

如 $\text{MSE}<T_s$，$T_s>0$是一个预定的阈值，则表示子块 R 处于静止区，$R(i,j,t)$完全可用 $R(i,j,t-1)$来代替，如果 $R(i,j,t-1)$已在自己帧内找到匹配父块，则 $R(i,j,t)$可用其 $R(i,j,t-1)$自仿射象素表示，只要几个比特的信息量。如果 $\text{MSE}>T_s$，表示该子块是一个运动块，同样采用块的匹配方式来找其相匹配的父块，只是不在本帧内寻找，而是到前一帧中去寻找。

利用(6.7)式，令$(t-1)$时刻的图象为 $I^2(t-1)$，t 时刻的图象为 $I^2(t)$，利用找块的原理，有

$$B_{i,j}^R I^2(t) \approx C_{i,j} L_{P(i,j)} A_V B_{\pi(k,l)}^D I^2(t-1) + h_{i,j} B_{i,j}^R I. \tag{8.7}$$

I 为全 1 的图象。其误差为

$$\begin{aligned}E_d = \min \| \sum_{ij} (B_{i,j}^R I^2(t) \\ - C_{i,j} L_{P(i,j)} A_V B_{\pi(k,l)}^D I^2(t-1) + h_{i,j} B_{i,j}^R I). \|.\end{aligned}$$

ii. 三维的分形编码方法

三维分形编码[39]是把一个图象序列看为一个整体，利用运动图象的前后帧之间存在的很大相关性，把 M 帧合为一个立体的三维块，子块和父块都从原来的二维变为三维；如果考虑在三维空间中的一个灰阶值 $u(i,j,t)$，i,j 为二维的平面上的坐标，t 为时间坐标，其取块的方式为 $B_{i,j,k}^R I^3$，对完备度量空间$(\mathscr{X},d)$中的两个立体块 μ、ν，都是具有相同的长、宽和时间的立方体，如其三维坐标上的灰阶值分别为 $\mu(i,j,k)$和 $\nu(i,j,k)$，其立方体块之间的匹配误差为

$$d(\mu,\nu) = \frac{1}{K \times L \times M}\sum_{i=1}^{K}\sum_{j=1}^{L}\sum_{k=1}^{M}(\mu(i,j,k)-\nu(i,j,k))^2,$$

$$\forall \mu,\nu \in \mathscr{X} \tag{8.8}$$

同样，也能利用四叉树的方法进行立方体之间的划块，"取块"是在

I^3的空间中进行的，记为$B_{K,L,M}^{*}I^3$. 在具体操作时，同样把I^3图象分为不相交的子块R_q；从父块到子块的变换为w_q；总的序列图象的变换为

$$W: \quad \mathscr{H}(\ ,\ ) \to \mathscr{H}(\ \ )$$

$$W(B) = \bigcup_{q=1} w_q \qquad w_q: D_q \to R_q$$

由于三维立方体的变换十分复杂，在立方体块内的每个点可用点对点的变换的形式来表示，每个变换w_q写为

$$w_q\begin{vmatrix} x \\ y \\ t \\ z \end{vmatrix} = \begin{vmatrix} a_q & b_q & e_q & 0 \\ c_q & d_q & f_q & 0 \\ g_q & h_q & 0_q & 0 \\ 0 & 0 & 0 & s_q \end{vmatrix}\begin{vmatrix} x \\ y \\ t \\ z \end{vmatrix} + \begin{vmatrix} \Delta x_q \\ \Delta y_q \\ \Delta t_q \\ p_q \end{vmatrix}. \tag{8.10}$$

这里$z=u(x,y,t)$是图象的灰阶函数. 在实际操作时，(8.10)式太复杂，可把它拆为空间变换和时间变换两部分. 对于空间变换T_q，有

$$T_q\begin{vmatrix} x \\ y \\ t \\ z \end{vmatrix} = \begin{vmatrix} a_q & b_q & 0 & 0 \\ c_q & d_q & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & s_q \end{vmatrix}\begin{vmatrix} x \\ y \\ t \\ z \end{vmatrix} + \begin{vmatrix} 0 \\ 0 \\ 0 \\ p_q \end{vmatrix}. \tag{8.11}$$

它作为一个三维的父块D_q到三维的子块R_q的映射，把D_q首先压缩$\frac{1}{4}$(这是经过四点平均而得到的)，然后改变D_q的灰阶值(通过加偏移P_q和尺度因子S_q而得到的)，使D_q和R_q得到最优匹配，而参数a_q, b_q, c_q, d_q表示位置的旋转因子，对于量化的图象往往采用对称和反射方式来作旋转，此时时间维t是不变的. 与二维方式相似，其变换见图 8-3.

空间域的变换明显是压缩变换. 此时，时间t不变. 在时间域的变换中，子块和父块往往是同一个尺寸，因为时间域中帧与帧之间的冗余本来是处于同尺度的块上. 为了不使多次迭代造成不收敛，在时间域上的变换t_q用如下形式：

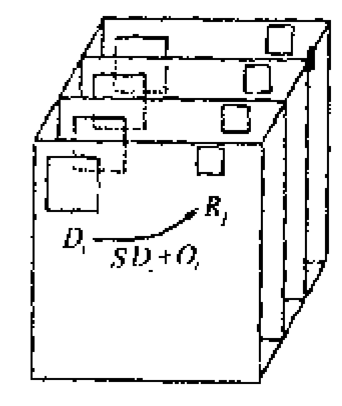
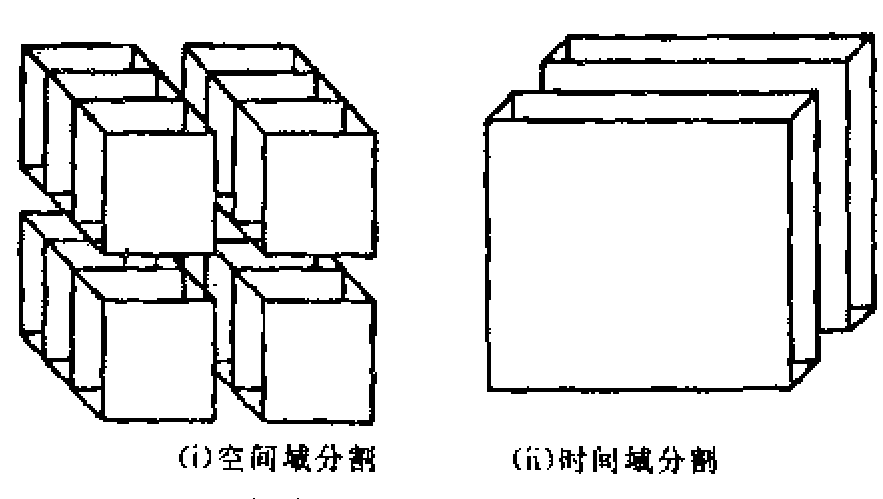


图 8-3 空间域的变换

$$t_q\begin{pmatrix}x\\y\\t\\z\end{pmatrix}=\begin{pmatrix}x\\y\\t\\z\end{pmatrix}+\begin{pmatrix}\Delta x\\\Delta y\\\Delta t\\0\end{pmatrix},\qquad(8.12)$$

这里 Δt 为负数，Δt 和 Δy 为平移参数.通过搜索，找到最好的 D_q 块与 R_q 进行匹配.事实上，运动模块 Δt 不能取得太大，一般为 -1、-2或 -3，因为每一帧的 D_q 可能有一些延迟，这样在 Δt 取得很大时，就很难找到其相应的父块.

四叉树分割的方法同样可用于三维的情况.对于空间分割，与二维情况相似；而对于时间方向的分割，可用二叉树方法.如图 8-4所示.



(i)空间域分割　　(ii)时间域分割

图 8-4 父、子块的分割情况

在什么时候进行分割，取决于下列公式

$$\text{分割}\begin{cases}\text{空间域,} & \text{如果}\sum_{i=1}^{4}E_d(W_i(D_i^s),R_i^s)<\sum_{j=1}^{2}E_d(W_j(D_j^t),R_j^t);\\ \text{时间域,} & \text{其他.}\end{cases}\qquad(8.13)$$

这里，上角的标号 s、t 分别表示已经过空间域分割和时间域分割的块.在搜索时，通过从顶端到底端的搜索，自动按照(8.13)式判别其分割，最后达到最优匹配的要求.

在具体操作时，还分为三种情况：

i. 子块与父块都十分平坦，而且没有运动。这种区域主要用空间变换分割，如块的尺寸比较大，编码效率很高。

ii. 子块与父块都十分复杂，但没有运动。这些块的编码可采用时间编码，因为立方体的编码其效率高，而且也比较省时间，可用空间与时间同时分割。

iii. 如果块比较平坦，而且有运动，这时也用时间编码分割的方式进行。

图 8-5表示序列图象压缩的结果。

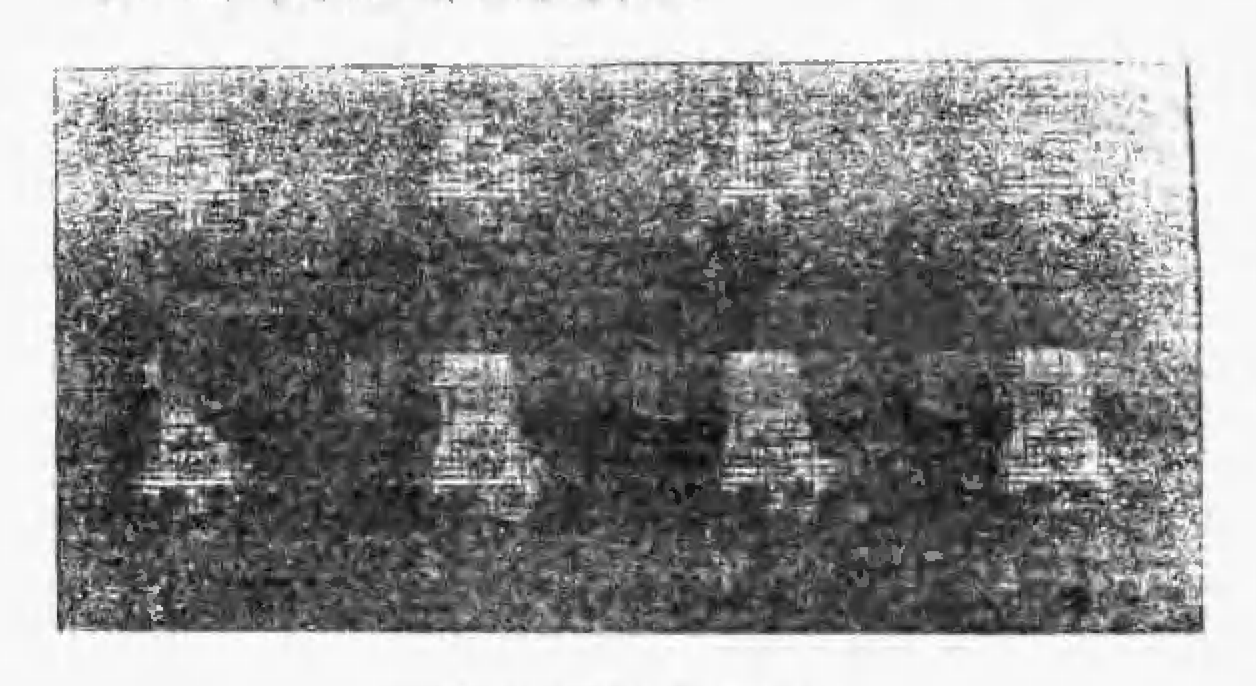


图8-5 运动图象的自适应分割

第3章

与其他方法结合的分形编码

分形编码是建立在非线性分形几何理论基础上的，它的特点是，任何有灰阶的图象都可以用一组迭代函数进行迭代而得到；压缩比比较高；解码的速度快. 但其编码速度，尤其是子块和父块匹配的搜索复杂度高，很难做到实时处理. 上一章中用了很多改进的方法来提高搜索速度，都是建立在分块的基础上. 像巴恩斯利那样严格采用分割的方法解逆问题，还需进一步探讨. 因此，分形编码还处于研究阶段，它尚未达到成熟. 另外，余弦变换法和小波变换法有它们各自的优点. 近年来，有关分形与其他方法的结合的研究正在开展. 下面介绍分形与小波变换结合的方法以及分形与余弦变换结合的方法.

§9 小波变换简介

小波分析和小波变换(wavelet transform)是近年来迅速发展起来的新兴科学，它同时具有理论和应用的意义. 它的分析方法类似于傅里叶分析，但与正弦和余弦基不同，而是一种新的表示函数的基. 小波变换又可与多分辨率结合起来，因此可采用不同分辨率的不同量化达到编码的目的.

1. 傅里叶分析和小波分析

设 $L^2(0,2\pi)$ 是在 $(0,2\pi)$ 上定义的所有可测且 $\int_0^{2\pi}|f(x)|^2dx < \infty$ 的集合. $L^2(0,2\pi)$ 可延拓到实数轴，即 $f(x)=f(x-2\pi)$，对一切 x 成立. $L^2(0,2\pi)$ 称为 2π 周期平方可积函数空间. 傅里叶变

换的本质，是将 $L^2(0,2\pi)$ 中的一个任意函数 $f(x)$ 表示为一族标准的正交函数 $\{e^{i\omega_n x} | n\in \mathbb{Z}\}$ 的加权求和表示式：

$$f(x) = \sum_{n=-\infty}^{\infty} c(\omega_n)e^{i\omega_n x}. \tag{9.1}$$

其中系数 $c(\omega_n)$ 满足

$$c(\omega_n) = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-i\omega_n x}\,dx, \tag{9.2}$$

为 $f(x)$ 的傅里叶系数．即 $f(x)$ 可分解为无限多个互相正交的分量 $g_n(x) = c(\omega_n)e^{i\omega_n x}$ 的和．所谓正交，是指内积

$$\langle g_m(x), g_n(x)\rangle = 0, \qquad \forall\, n \neq m.$$

$$\langle g_m(x), g_n(x)\rangle = \frac{1}{2n}\int_0^{2\pi} g_m(x)\,\overline{g_n}(x)dx. \tag{9.3}$$

在傅里叶变换中，基函数是 $\sin x$ 和 $\cos x$ 以及他们的多次谐波．由于 $f(x)$ 和它的傅里叶变换系数具有完美的对称形式，我们可将对原来函数 $f(x)$ 的研究转化为对权系数的研究．经典的傅里叶分析是一种纯频域分析．

傅里叶变换有其固有的缺点：即在任何有限频段上的信息，都不足以确定在任意小范围内的函数 $f(x)$，有界的函数 $f(x)$ 却要用无限的基函数来迭加，时(空)间域中的有界带来了频率域上的无限．这是因为傅里叶变换的基函数是在时(空)域上无限的．

是否能找到一个正交基，它在时(空)域和频域中都是比较集中．这就可以解决在傅里叶变换中的问题，达到时域频域局部分析的目的．小波变换就是在此基础上提出的．

如同 $L^2(0,2\pi)$ 空间中的傅里叶变换基一样，令 $L^2(\mathbb{R})$ 是实数轴上可测函数组成的平方可积空间，如果找到一个函数 $\psi(x)$，满足

$$C_\psi = 2\pi\int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|}d\omega < \infty, \tag{9.4}$$

$$\int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty. \tag{9.5}$$

这里 $\psi(\omega)$是 $\psi(x)$的 频 率 域 表 示 . 并且,在 $L^2(\mathbb{R})$上的任意函数 $f(x)$,满足

$$f(a,b) = \frac{1}{\sqrt{c_h}} \cdot \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(x) \bar{\psi}\left(\frac{x-b}{a}\right) dx, \tag{9.6}$$

$$f(x) = \frac{1}{\sqrt{c_h}} \int f(a,b) |a|^{-\frac{1}{2}} \psi\left(\frac{x-b}{a}\right) \left(\frac{dadb}{a^2}\right), \tag{9.7}$$

则 $\psi(x)$为其小波函数.

从(9.4)、(9.5)式看到,$\psi(x)$在时域和频域上都是有界的,在 $\pm\infty$时衰减到 0. 为了让 $\psi(x)$离散化,且生成整个 $L^2(\mathbb{R})$空间,取 $a = a_0^m (a_0 > 1)$, $b = nb_0 a_0^m (b_0 \in \mathbb{R})$, 则

$$\left\{\psi_{m,n}(x) = |a_0|^{-\frac{m}{2}} \psi(a_0^{-m} x - nb_0): m,n \in \mathbb{Z}\right\} \tag{9.8}$$

能构成 $L^2(\mathbb{R})$上的规范正交基,且有良好的时、频域局部性.

式(9.8)中的 nb_0 表示小波函数 ψ 在实轴 x 上的平移,当 n 趋向$\pm\infty$时,ψ 可覆盖在整个实空间. a_0^{-m}表示非单一频率的波,称膨胀或缩小因子,正交基为

$$\psi(a_0^{-m} x - nb_0). \qquad m,n \in \mathbb{Z}$$

考虑到 $\psi \in L^2(\mathbb{R})$具有规范的形式,$\|\psi_{m,n}\| = 1$,则

$$\psi_{m,n}(x) = |a_0|^{-\frac{m}{2}} \psi(a_0^{-m} x - nb_0). \qquad m,n \in \mathbb{Z}$$

当 $a_0 = 2$ 时,该小波基称为二进小波基.

定义 1 如果$\{\psi_{m,n}\}$是一个规范正交基,即

$$\langle \psi_{m,n}, \psi_{j,k} \rangle = \delta_{mj} \delta_{nk}, \qquad j,k,m,n \in \mathbb{Z} \tag{9.9}$$

这里

$$\delta_{mj} = \begin{cases} 1, & \text{当 } m = j \text{ 时;} \\ 0, & \text{当 } m \neq j \text{ 时,} \end{cases}$$

$$\delta_{nk}=\begin{cases}1, & \text{当 } n=k \text{ 时;}\\ 0, & \text{当 } n\neq k \text{ 时,}\end{cases}$$

则称函数 $\psi\in L^2(\mathbb{R})$是正交小波.

可以证明,任意函数 $f\in L^2(\mathbb{R})$ 能写成

$$f(x)=\sum_{m,n\in\mathbb{Z}}^{\infty} C_{m,n}\psi_{m,n}(x); \tag{9.10}$$

其变换系数 $C_{m,n}$满足

$$C_{m,n}=\langle f,\psi_{m,n}\rangle. \tag{9.11}$$

对照(9.1)、(9.2),离散小波变换与离散的傅里叶变换有相同的形式.

小波基函数的另一个要求为

$$\int_{-\infty}^{\infty}\psi(x)dx=0. \tag{9.12}$$

考虑到条件(9.4)、(9.5)、(9.12),不同尺度、不同平移因子的三个小波基如图 9-1 所示.



图 9-1　不同膨胀因子、平移因子的小波基

图 9-1 中,最左面的小波基的膨胀度最大,代表比较低的频率,是坐标右移而得到的;最右面的小波基膨胀度最小,代表高频率的情况,是坐标右移而得到的.

虽然,小波变换的基在一般情况下需要完全正交,这是由于人们希望其变换是可逆的.但是从函数重构的情况看,现在发展为一种双正交的小波基,在函数正变换时用小波基 $\psi^{ab}(x)$,而在逆变换时采用另一组小波基 $\tilde{\psi}^{ab}(x)$, $a=a_0^{\,m}$、$b=nb_0a_0^{\,m}$. $\tilde{\psi}^{ab}(x)$ 称为

$\psi^{ab}(x)$ 的对偶小波基.

对于任意函数 $f(x) \in L^2(\Re)$，其正变换和逆变换分别为：

$$(W_\psi f)(a,b) = \langle f(x),\ \psi^{ab}(x)\rangle, \tag{9.13}$$

$$f(x) = \langle (W_\psi f)(a,b),\ \tilde{\psi}^{ab}(x)\rangle. \tag{9.14}$$

对于正交的小波基函数，$\tilde{\psi}^{ab}(x) = \psi^{ab}(x)$. $\tilde{\psi}^{ab}(x)$ 是一个重构的小波基. 一般讲，“对偶”不一定是唯一的.

小波变换与傅里叶变换的最大区别是：

i. 傅里叶变换得到的频谱需要无限的时间信息，如果在时间信息上加窗口，则在频率域上变成无限的，从而不能做到时域和频域的局域性. 小波变换的函数本身在时域上是局域的，在设计中又考虑频域上的局域性，因而它被称为时-频分析的新工具.

ii. 在傅里叶变换中，对于窗口的时间间隔是固定的，但实际分析信号时，希望高频谱的信息，时间间隔要相对小，才能给出比较好的精度；对于低频谱的信息，时间间隔要相对地宽，以给出完全的信息. 当在高低频都存在的信号中，很难用傅里叶变换进行分析. 小波变换的不同尺度是使时间-频率窗十分灵活，在高中心频率时可自动变窄而在低中心频率时自动变宽. 从信号分析的角度，小波变换优于傅里叶变换.

2. 分辨率分析

马莱德(S. E. Mallat)[41]在 1989 年提出了一种小波分析的方法，即多分辨率分析. 如 § 7 中提到的多分辨率方法表示一样，如果这里的分辨率也以 2 的幂次变化，分辨率变低是用四点平均的方法. 为了简单起见，先讨论一维的情况，让 A_{2^j} 算子表示为信号以 2^j 分辨率上的近似算子，例如原始信号 $f(x) \in L^2(\Re)$，则算子 A_{2^j} 具有下述性质：

i. A_{2^j} 是一个线性算子. 假定 $A_{2^j}f(x)$ 是 $f(x)$ 在 2^j 分辨率上的近似，则 $A_{2^j} \circ A_{2^j} = A_{2^j}$，$A_{2^j}$ 是在一个矢量空间 V_{2^j} 上，$V_{2^j} \subset L^2(\Re)$，V_{2^j} 可解释为在 2^j 的分辨率上所有的可能的近似.

ii. 在 2^j 分辨率上所有近似函数中，$A_{2^j}f(x)$ 最近似于 $f(x)$：

$$\| g(x) - f(x) \| \geqslant \| A_{2^j} f(x) - f(x) \|,$$

$$\forall g(x) \in V_{2^j} \quad (9.15)$$

因此，A_{2^j}是在矢量空间 V_{2^j}上的垂直投影算子.

iii. 信号在 2^{j+1}分辨率上的近似包含了所有的在 2^j 上的近似，即

$$V_{2^j} \subset V_{2^{j+1}}, \qquad \forall j \in \mathbb{Z} \quad (9.16)$$

iv. 在近似函数空间，从一个分辨率到另一个分辨率的变化为

$$f(x) \in V_{2^j} \Leftrightarrow f(2x) \in V_{2^{j+1}},$$

$$\forall j \in \mathbb{Z} \quad (9.17)$$

v. 信号 $f(x)$在 2^j 分辨率的近似 $A_{2^j} f(x)$，可用每单位长度 2^j 个采样点来离散.

vi. 不同分辨率的闭空间 V_j，满足：

$$\lim_{j \to +\infty} V_{2^j} = \bigcup_{j=-\infty}^{\infty} V_{2^j} = L^2(\mathbb{R}), \quad (9.18)$$

$$\lim_{j \to -\infty} V_{2^j} = \bigcap_{j=-\infty}^{\infty} V_{2^j} = \{0\}. \quad (9.19)$$

对于已知的原始函数 $f(x)$，我们假设它在 $j=0$ 的空间中，其分辨率为最高，而其他分辨率为 $2^j, j=-1,-2,-3,\cdots\cdots$为不同的分辨率空间. 对 $j=0$ 空间中存在一个 $\varphi \in V_{2^0}$，使$\{\varphi_{0,k}, k \in \mathbb{Z}\}$是 V_{2^0}的正交归一基. 按照上述性质，有$\{\varphi_{j,k}, k \in \mathbb{Z}\}$也是 V_{2^j}空间的正交基，φ称为尺度函数. 下面给出 φ的形式和理论.

定理 1 令$(V_{2_j})_{j \in \mathbb{Z}}$是 $L^2(\mathbb{R})$上的多分辨率近似，存在唯一的函数 $\varphi(x) \in L^2(\mathbb{R})$，称为尺度函数，如果设 $\varphi_{2^j}(x) = 2^{j/2} \varphi(2^j x), j \in \mathbb{Z}$，则 $\sqrt{2^{-j}} \varphi_{2^j}(x - 2^{-j}k)_{k \in \mathbb{Z}}$是 V_{2^j}空间中的正交归一基.

证明见[41].

现在我们把一个原始信号 $f(x)$通过尺度函数投影到 V_{2^j}空间中去：

$$A_{2^j} f(x) = 2^{-j} \sum_{k=-\infty}^{\infty} \langle f(u), \varphi_{2^j}(u - 2^{-j}k) \rangle \varphi_{2^j}(x - 2^{-j}k),$$

$$\forall f(x) \in L^2(\mathbb{R}) \quad (9.20)$$

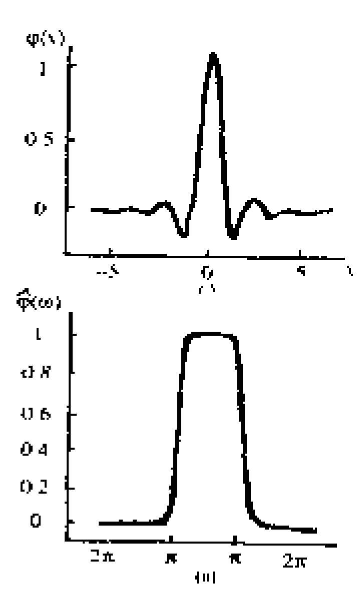


图 9-2 尺度函数

图 9-2 表示尺度函数 $\varphi(x)$ 和它的频率域表示的 $\hat{\varphi}(\omega)$. 可以看到,尺度函数具有低通滤波的作用,在 2^j 分辨率上的近似可看作是信号与该分辨率尺度函数的内积,或对于一个离散信号来说可为对采样率为 2^j 的信号的空间的滤波.

如把尺度函数写成离散的形式,并与小波基的形式对照起来,

$$\varphi_{j,k}(x) := 2^{j/2}\varphi(2^j x - k).$$

$$j,k \in Z \qquad (9.21)$$

其中 2^j 为膨胀因子,犹如小波基中的频率变化. 如果 $\varphi(x) \in V_{2^j}$,则 $\varphi(2x) \in V_{2^{j+1}}$(性质 3),则 $\varphi(x)$ 可以用 $\varphi(2x-k)$ 来表示. 对于二进小波,存在一个唯一的序列 $h(k)$,可以描述两个尺度函数之间的关系(称为双尺度方程)

$$\varphi(x) = \sum_{k=-\infty}^{\infty} h(k)\varphi(2x-k). \qquad (9.22)$$

$h(k)$ 同样有低通滤波的作用. 因此,不同的尺度函数对函数 $f(x)$ 的卷积,可以化为 $h(k)$ 对函数的滤波后二点取一点而得到. 例如 $f(x)$ 被空间采样为 2^j 个点,则可认为 δ 函数就是分辨率最高的 φ 函数,因此 $A_{2^j}f(x)=f(x)$. 如果将 $f(x)$ 经 $h(k)$ 的滤波器滤波,则相当于 $f(x)$ 与 $V_{2^{j-1}}$ 空间中的尺度函数卷积. 如此时取样为 2^{j-1} 点,则可得到在 $V_{2^{j-1}}$ 空间中的近似 $A_{2^{j-1}}f(x)$. 如继续用 $h(k)$ 的低通滤波器滤波,并二点取一点, 可得 $A_{2^{j-2}}f(x)$, $A_{2^{j-3}}f(x)$, 直到 $f(x)$ 只剩下一点为止,则就可得到任意分辨率的近似.

现在考虑函数 $f(x)$ 在 2^j 和 2^{j-1} 分辨率上近似的差值. 我们称它为误差信号. 它是在 2^j 分辨率的空间中,考虑在 V_{2^j} 空间中与 $V_{2^{j-1}}$ 正交的补空间,误差信号是投影到此子空间中,让 $O_{2^{j-1}}$ 表示在 V_{2^j} 中与 $V_{2^{j-1}}$ 正交的补空间,有

$$V_{2^j} = V_{2^{j-1}} + O_{2^{j-1}}, \qquad (9.23)$$

$$V_{2^j} \cap O_{2^{j-1}} = \{0\}. \tag{9.24}$$

可以证明，在 $O_{2^{j-1}}$ 空间上的基函数亦能够用尺度函数来建立，就是小波基函数 $\psi(x)$. 由於 $\psi(x)$ 也是属于上一个尺度空间的函数，可参照(9.22)，如果 $\psi(2x) \in V_{2^j}$，$\varphi(2x) \in V_{2^{j+1}}$，$\psi(x)$ 也可用 $\varphi(2x-k)$ 来表示，有

$$\psi(x) = \sum_{k=-\infty}^{\infty} g(k)\,\varphi(2x-k). \tag{9.25}$$

$g(k)$ 是一个高通滤波器，作为 $h(k)$ 低通的补，对于 $f(x)$ 与小波函数的内积，也可像尺度函数一样，用高通滤波器滤波而得到. 例如在最高分辨率的近似为 $A_{2^j} f(x) = f(x)$，如将 $f(x)$ 经 $g(k)$ 滤波，并二点取一点，则得到为 $V_{2^{j-1}}$ 空间中的小波系数. 如在 $A_{2^{j-1}} f(x)$ 的空间中进行 $g(k)$ 滤波，又得到 $V_{2^{j-2}}$ 空间中的小波系数. 这样，小波变换不是采用不同尺度的小波基进行卷积，而只要用 $h(k)$ 和 $g(k)$ 两个滤波器进行滤波就可得到有关的小波变换系数. 当然，实际上的滤波器都是 FIR 滤波器，从而(9.22)和(9.25)可改为：

$$\varphi(x) = \sum_{k=1}^{n} h(k)\varphi(2x-k). \tag{9.26}$$

$$\psi(x) = \sum_{k=1}^{m} g(k)\,\varphi(2x-k). \tag{9.27}$$

定理 2　设 $(V_{2^j})_{j\in Z}$ 是一个多分辨率的矢量空间，$\varphi(x)$ 是其尺度函数. 将(9.27)、(9.26)化为傅里叶变换的频率域，令 $H(\omega)$ 和 $G(\omega)$ 为低高通滤波在频率域表示，则在频率域 $\hat{\psi}(\omega)$，$\hat{\varphi}(\omega)$，满足

$$\hat{\psi}(\omega) = G\left(\frac{\omega}{2}\right)\hat{\varphi}\left(\frac{\omega}{2}\right),$$

$$\hat{\varphi}(\omega) = H\left(\frac{\omega}{2}\right)\hat{\varphi}\left(\frac{\omega}{2}\right),$$

且

$$G(\omega) = e^{-i\omega}\overline{H(\omega+\pi)}. \tag{9.28}$$

令

$$\psi_{2^j}(x) = 2^j\psi(2^j x)$$

表示 $\psi(x)$ 被尺度因子 2^j 膨胀，那么 $(\sqrt{2^{-j}}\psi_{2^j}(x-2^{-j}k))_{k\in\mathbb{Z}}$ 是 O_{2^j} 的正交归一基，且有 $(\sqrt{2^{-j}}\psi_{2^j}(x-2^{-j}k))_{j,k\in\mathbb{Z}}$ 是 $L^2(\mathbb{R})$ 的一个正交归一基. $\psi(x)$ 被称为正交小波.

证明见[40].

式(9.28)告诉我们，只要已知低通滤波器 $H(\omega)$，就可以得到高通滤波器 $G(\omega)$.(9.28)在时间域上的关系为

$$g(k) = (-1)^{1-k}h(1-k). \tag{9.29}$$

多分辨率的方法不仅给出了用滤波器求小波变换的方法，同时也给小波基的设计提供了一个途径.

3. 小波基的种类和构造

二进小波基是最常用的小波基，它的发展最早. 近年来，多进的小波基也发展起来. 这里只对与分形结合的一些小波基做一粗略的介绍.

i. Haar 小波

取 $V_{2^0}=\{u \mid u\in L^2(\mathbb{R})$，在每一区间 $(k,k+1)$ 上，$u(x)=$ 常数$\}$，函数 $\varphi(x)=\chi_{(0,1)}$；$\chi_{01}(x)=\begin{cases}1, & \text{当 } x\in[0,1] \text{ 时；}\\ 0, & \text{其他，}\end{cases}$ $\varphi(x)$ 所形成的整数平移系 $\{\varphi(x-k) \mid k\in\mathbb{Z}\}$ 是闭子空间 V_{2^0} 的规范正交基：

$$\hat{\varphi}(\omega) = \frac{1}{\sqrt{2\pi}}e^{-i\omega/2}\frac{\sin(\omega/2)}{\omega/2},$$

它构成的多分辨率分析，是用尺度膨胀因子 2^j 构成，且 $\bigcup V_{2^j}=L^2(\mathbb{R})$. 利用式(9.26)、(9.27)、(9.28)，可得

$$H(\omega) = \frac{1}{2}(1+e^{i\omega}),$$

$$G(\omega) = e^{-i\omega}\overline{H(\omega+\pi)} = \frac{1}{2}(-1+e^{-i\omega}). \tag{9.30}$$

$$h(k) = \begin{cases} \dfrac{1}{\sqrt{2}}, & \text{当 } k = 0,1 \text{ 时;} \\ 0, & k \text{ 为其他时.} \end{cases}$$

$$g(k) = \begin{cases} -\dfrac{1}{\sqrt{2}}, & \text{当 } k = 0 \text{ 时;} \\ \dfrac{1}{\sqrt{2}}, & \text{当 } k = 1 \text{ 时;} \\ 0, & k \text{ 为其他时.} \end{cases} \tag{9.31}$$

Haar 小波是有限支撑的，而且比较简单，从而应用广泛. 但是，由于它在频率域是无限的，因而并不是最理想的小波基.

ii. 频率域是矩形的情况下的小波基

取 $V_{2^0} = \{u \in L^2(\mathbb{R}) \mid \sup p(u) \subset [-\pi, \pi]\}$，$V_{2^j}$ 具有平移不变性，在闭空间 V_{2^j} 上，有

$$V_{2^j} = \{u \in L^2(\mathbb{R}) \mid \sup p(u) \subset [-2^{-j}\pi, 2^{-j}\pi]\}.$$

$$\varphi(x) = \frac{1}{\sqrt{2\pi}}\chi_{(-\pi,\pi)}(\omega) = \frac{\sin x}{\pi x}. \tag{9.32}$$

利用(9.28)，$H\left(\frac{\omega}{2}\right) = \frac{\varphi(\omega)}{\varphi(\omega/2)}$，可得到

$$H(\omega) = \begin{cases} 1, & \text{当 } |\omega| < \dfrac{\pi}{2} \text{ 时;} \\ 0, & \text{当 } \dfrac{\pi}{2} < |\omega| < \pi \text{ 时,} \end{cases}$$

$$h(k) = \begin{cases} \dfrac{1}{\sqrt{2}}, & \text{当 } k = 0 \text{ 时;} \\ 0, & \text{当 } k = 2n \text{ 时;} \\ \dfrac{\sqrt{2}}{\pi}\dfrac{(-1)^n}{2n+1}, & \text{当 } k = 2n+1 \text{ 时.} \end{cases} \tag{9.33}$$

则有

$$G(\omega)=\begin{cases}e^{-i\omega}, & \text{当} -\dfrac{3\pi}{2}<\omega<-\dfrac{\pi}{2}\text{ 时;}\\ 0, & \text{当}\dfrac{\pi}{2}<\omega<0\text{ 或 }-2\pi<\omega<-\dfrac{3}{2}\omega\text{ 时.}\end{cases} \tag{9.34}$$

此小波在空间域中不是有限支撑的.

iii. Meyer 取非负偶数 $Q(\omega)\in C_0^{\infty}$, 满足条件:

① $\theta(\omega)=1$, 当 $0\leqslant\omega\leqslant 2\pi/3$ 时;

② $\theta(\omega)=0$, 当 $\omega\geqslant 4\pi/3$ 时;

③ $0<\theta(\omega)<1$, 当 $2\pi/3<\omega<4\pi/3$ 时;

④ $\theta^2(\omega)+\theta^2(2\pi-\omega)=1$.

令 $\varphi(x)=\dfrac{1}{\sqrt{2\pi}}\theta(x)$, 则 $\{\varphi(x-k)\,|\,k\in\mathbb{Z}\}$ 构成一个规范正交系. 可以计算, 得

$$H(\omega)=\theta(2\omega),$$

$$G(\omega)=e^{-i\omega}\theta(2\pi-2|\omega|).\qquad |\omega|\leqslant\pi \tag{9.35}$$

iv. 利用样条函数形成非正交的多分辨率的尺度函数, 并通过正交化方法得到 $\varphi(\omega)$、$H(\omega)$ 和 $G(\omega)$. 设 V_{2^0} 是所有整单位 $[k,k+1)$ 上的二阶连续 $L^2(\mathbb{R})$ 函数构成的闭子空间. 对每个整数 k, 闭子空间 V_k 可按二进伸缩 V_0 得到. $\{V_k\}_{k\in\mathbb{Z}}$ 构成 $L^2(\mathbb{R})$ 的一个多分辨分析. 用 4 阶样条得到非正交函数为

$$f=\chi_{[0,1]}*\chi_{[0,1]}*\chi_{[0,1]}*\chi_{[0,1]},$$

然后根据正交化方法, 得到

$$\hat{\varphi}(\omega)=\frac{1}{\sqrt{2\pi}}\,\frac{1}{\omega^2\sqrt{\Sigma_8(\omega)}}, \tag{9.36}$$

$$H(\omega)=\sqrt{\frac{\Sigma_8(\omega)}{2^8\Sigma_8(2\omega)}}, \tag{9.37}$$

$\Sigma_2(\omega)=\dfrac{1}{4\sin^2\left(\dfrac{\omega}{2}\right)}$, $\Sigma_8(\omega)$ 是通过计算 $\Sigma_2(\omega)$ 的 6 阶导数而得,

$$G(\omega) = e^{-i\omega}\overline{H(\omega+\pi)}. \tag{9.38}$$

v. 有限正交小波

上面几种小波都是用多分辨率的分析方法首先设计尺度函数然后得到低通滤波器和高通滤波器. I. Daubechies[40]首先给出频率域的滤波器形式，是在正交有界条件下通过解多项式方程而得到的. 如果假设低通滤波器的形式为

$$H(e^{-i\omega}) = \left(\frac{1+e^{-i\omega}}{2}\right)^N P(e^{-i\omega}), \tag{9.39}$$

P 是一个实多项式，(9.39)展开后就是一个有限项的$(e^{-i\omega})$的多项式. 因此，如果 P 能解出来，则低通滤波器就可得到，且是有限项. 而 P 的多项式又要满足尺度函数和小波基函数的正交性等条件，则有

$$|H(e^{-i\omega})|^2 + |H(e^{-i\omega}+\pi)|^2 = 1. \tag{9.40}$$

而

$$|H(e^{-i\omega})|^2 = [\cos^2(\omega/2)]^N |P(e^{-i\omega})|^2. \tag{9.41}$$

由 $|P(e^{-i\omega})|^2 = \overline{P(e^{-i\omega})}p(e^{-i\omega})$ 可知，$|P(e^{-i\omega})|^2$ 能写成关于 $\cos\omega$ 的多项式，也可写成 $\cos^2\left(\frac{\omega}{2}\right)$ 或 $\sin^2\left(\frac{\omega}{2}\right)$ 的多项式. 若设 $y = \cos^2\left(\frac{\omega}{2}\right)$ $|P(e^{-i\omega})|^2 = Q(\sin^2\omega/L)$，则(9.40)可写为

$$y^N Q(1-y) + (1-y)^N Q(y) = 1. \tag{9.42}$$

且要满足

$$Q(y) \geqslant 0. \qquad (y \in [0,1])$$

通过数值计算，可求得 Q 多项式，然后求出 $H(e^{-i\omega})$，可得到紧支撑有界的小波基.

vi. 双正交小波(二进)

在(9.13)和(9.14)式中，我们给出了小波正变换和逆变换的公式，其中 $\tilde{\psi}_{jk}$为对偶小波基. 如小波基是正交的，则 $\tilde{\psi}_{jk}=\psi_{jk}$，$\tilde{\psi}_{jk}$是与 ψ_{jk}不同但有关联的小波基. 它定义尺度函数和小波函数 φ、ψ，及它们的对偶 $\tilde{\varphi}$、$\tilde{\psi}$，它们各自满足双尺度方程(9.26)和(9.27)：

$$\varphi(x)=\sum_k h(k)\varphi(2x-k)\rightarrow\hat{\varphi}(\omega)=H\left(\frac{\omega}{2}\right)\hat{\varphi}\left(\frac{\omega}{2}\right),\qquad(9.43)$$

$$\psi(x)=\sum_k g(k)\varphi(2x-k)\rightarrow\hat{\psi}(\omega)=G\left(\frac{\omega}{2}\right)\hat{\varphi}\left(\frac{\omega}{2}\right),\qquad(9.44)$$

$$\tilde{\varphi}(x)=\sum_{k'}\tilde{h}(k')\tilde{\varphi}(2x-k')\rightarrow\hat{\tilde{\varphi}}(\omega)=\tilde{H}\left(\frac{\omega}{2}\right)\hat{\tilde{\varphi}}\left(\frac{\omega}{2}\right),\qquad(9.45)$$

$$\tilde{\psi}(x)=\sum_{k'}\tilde{g}(k')\tilde{\varphi}(2x-k')\rightarrow\hat{\tilde{\psi}}(\omega)=\tilde{G}\left(\frac{\omega}{2}\right)\hat{\tilde{\varphi}}\left(\frac{\omega}{2}\right).\qquad(9.46)$$

多分辨率的形式也为

$$V_{-2}\subset V_{-1}\subset V_0\subset V_1\cdots\cdots,$$
$$\tilde{V}_{-2}\subset\tilde{V}_{-1}\subset\tilde{V}_0\subset\tilde{V}_1\cdots\cdots,$$

且有正交条件为

$$O_j\perp\tilde{V}_j;\quad\tilde{O}_j\perp V_j.$$

按照重构条件及双正交条件，要满足下列三个方程：

重构条件：

$$\tilde{h}(e^{-j\omega})\cdot\overline{h(e^{-j\omega})}+\tilde{g}(e^{-j\omega})\cdot\overline{g(e^{-j\omega})}=2,$$

$$\tilde{h}(e^{-j\omega})\cdot\overline{h(-e^{-j\omega})}+\tilde{g}(e^{-j\omega})\cdot\overline{g(-e^{-j\omega})}=0;$$

正交条件：

$$H(\omega)\cdot\overline{\tilde{H}(\omega)}+H(\omega+\pi)\cdot\overline{\tilde{H}(\omega+\pi)}=1.$$

仍按 v 中的方法解得双正交小波基，即 $H(\omega)$、$\tilde{H}(\omega)$、$G(\omega)$、$\tilde{G}(\omega)$. 双正交小波可满足紧支撑、对称等条件，十分灵活.

vii. 多通道小波

上面讨论的小波基函数都是二进小波. 多进小波是建立在多通道滤波器的基础上设计多通道滤波器，再用上述双正交的概念来获得多通道的小波. 本书中，分形与小波结合的方法中我们只用了二进小波，因此这里就不多叙述.

4. 小波变换对一维函数的分解和重构

多分辨分析中，对 $f(x)\in L^2(\mathbb{R})$ 在 2^j 分辨率上的近似，为

$$A_{2^j}f(x)=2^{-j}\sum_k\langle f(u),\varphi_{2^j}(u-2^{-j}k)\rangle\varphi_{2^j}(x-2^{-j}k).$$

如果对在 V_{2^j} 空间中的尺度函数 $\varphi_{2^j}(x-2^jk)$ 用投影到 $V_{2^{j+1}}$ 空间中来表示：

$$\varphi_{2^j}(x-2^{-j}n)=2^{-j-1}\sum_{k=-\infty}^{\infty}\langle\varphi_{2^j}(u-2^{-j}n),\ \varphi_{2^{j+1}}(u-2^{-j-1}k)\rangle\varphi_{2^{j+1}}(x-2^{-j-1}k), \tag{9.51}$$

令 $u'=(u-2^{-j}n)$，变换后得到

$$2^{-j-1}\langle\varphi_{2^j}(u-2^{-j}n),\ \varphi_{2^{j+1}}(u-2^{-j-1}k)\rangle=\langle\varphi_{2^{-1}}(u),\ \varphi(u-(k-2n))\rangle.$$

在(9.51)的两边对 $f(x)$ 求内积，有

$$\langle f(u),\varphi_{2^j}(u-2^{-j}n)\rangle=\sum_k\langle\varphi_{2^{-1}}(u),\varphi(u-(k-2n))\rangle\cdot\langle f(u),\varphi_{2^{j+1}}(u-2^{-j-1}k)\rangle.$$

根据双尺度方程(9.22)，

$$h(n)=\langle\varphi_{2^{-1}}(u),\ \varphi(u-n)\rangle.$$

所以

$$\langle f(u),\varphi_{2^{-1}}(u-2^{-j}n)\rangle=\sum_k h(2n-k)\cdot\langle f(u),\varphi_{2^{j+1}}(u-2^{-j-1}k)\rangle. \tag{9.52}$$

因此，$f(x)$ 在 2^j 分辨率上的近似，相当于滤波器 h 在 2^{j+1} 分辨率上滤波后二点取一点的结果.

同样，对于 2^j 分辨率上的小波函数 $\psi_{2^j}(x-2^{-j}n)$，是建立在 O_{2^j} 上，$O_{2^j}\subset V_{2^{j+1}}$，也可以用 $V_{2^{j+1}}$ 上的基来表示：

$$\psi_{2^j}(x-2^{-j}n)=2^{-j-1}\sum_k\langle\psi_{2^j}(u-2^{-j}n),\ \varphi_{2^{j+1}}(u-2^{-j-1}k)\rangle\varphi_{2^{j+1}}(x-2^{-j-1}k). \tag{9.53 a}$$

同理，有

$$2^{-j-1}\sum_k\langle\psi_{2^j}(u-2^{-j}n),\varphi_{2^{j+1}}(u-2^{-j-1}k)\rangle=\langle\psi_{2^{-1}}(u),\varphi(u-(k-2n))\rangle. \tag{9.53 b}$$

(9.53)两边对 $f(x)$ 求内积，得

$$\langle f(u),\psi_{2^j}(u-2^{-j}n)\rangle = \sum_k \langle \psi_{2^{-1}}(u),\varphi(u-(k-2n))\rangle \cdot \langle f(u),\varphi_{2^{j+1}}(u-2^{-j-1}k)\rangle. \quad (9.54)$$

根据双尺度方程(9.27)，有

$$g(n) = \langle \psi_{2^{-1}}(u),\ \varphi(u-n)\rangle.$$

所以，(9.54)可写为

$$\langle f(u),\psi_{2^j}(u-2^{-j}n)\rangle = \sum_k g(2n-k)\langle f(u),\varphi_{2^{j+1}}(u-2^{-j-1}k)\rangle. \quad (9.55)$$

从(9.52)和(9.55)式，一个 2^{j+1} 分辨率的函数可近似分解，如图9-3所示.

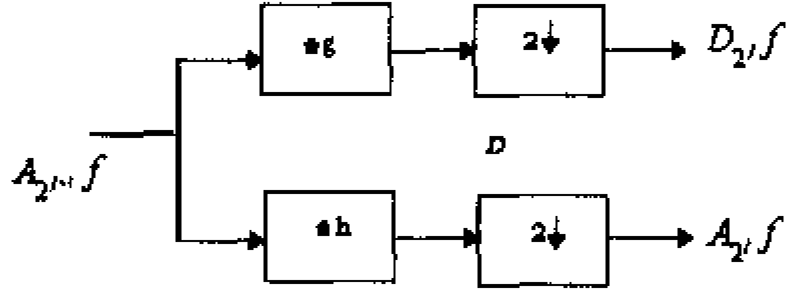


图 9-3 小波分解

图9-3中的 $D_{2^j}f$，是 $f(x)$ 在小波 ψ_{2^j} 上的系数. $A_{2^j}f$ 还可以继续分解，得到 $D_{2^{j-1}}f, D_{2^{j-2}}f$，……从而得到所有的在不同膨胀系数下的小波系数.

由于 $V_{2^{j+1}}=V_{2^j}+O_{2^j}$，$V_{2^j}\perp O_{2^j}$，从而重构是建立在 $V_{2^{j+1}}$ 中的正交规范基 $\varphi_{2^j}(x-2^{-j}n)$ 和 $\psi_{2^j}(x-2^{-j}n)$ 上. 其关系满足

$$\varphi_{2^{j+1}}(x-2^{-j-1}n) = 2^{-j}\sum_k \langle \varphi_{2^j}(u-2^{-j}k), \varphi_{2^{j+1}}(u-2^{-j-1}n)\rangle \varphi_{2^j}(x-2^{-j}k) + 2^{-j}\sum_k \langle \psi_{2^j}(u-2^{-j}k), \varphi_{2^{j+1}}(u-2^{-j-1}n)\rangle \psi_{2^j}(x-2^{-j}k). \quad (9.56)$$

(9.56)两边对 $f(x)$ 求内积，可得

$$\langle f(u),\varphi_{2^{j+1}}(u-2^{-j-1}n)\rangle$$

$$= 2^{-j}\sum_k \langle \varphi_{2^j}(u - 2^{-j}k), \varphi_{2^{j+1}}(u - 2^{-j-1}n)\rangle$$

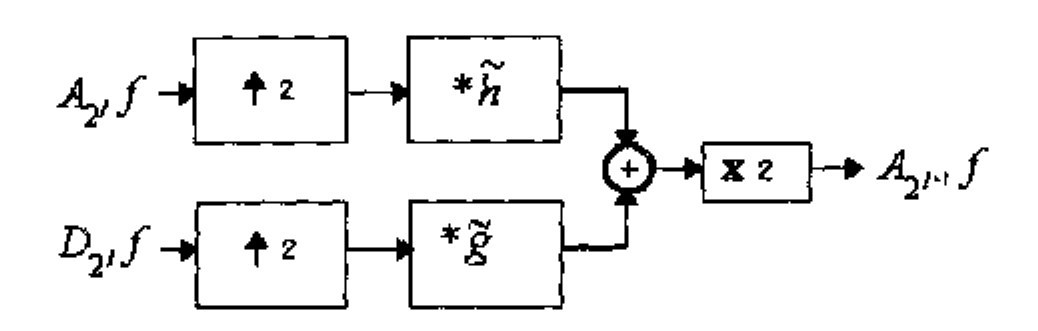
$$\cdot \langle f(u), \varphi_{2^j}(u - 2^{-j}k)\rangle + 2^{-j}\sum_k \langle \psi_{2^j}(u - 2^{-j}k),$$

$$\varphi_{2^{j+1}}(u - 2^{-j-1}u)\rangle \cdot \langle f(u), \psi_{2^j}(u - 2^{-j}k)\rangle$$

$$= 2\sum_k \tilde{h}(n - 2k)\langle f(u), \varphi_{2^j}(u - 2^{-j}k)\rangle$$

$$+ 2\sum_k \tilde{g}(n - 2k)\langle f(u), \psi_{2^j}(u - 2^{-j}k)\rangle. \qquad (9.57)$$

重建的框图如图 9-4 所示. 对正交小波, $\tilde{h}=h$, $\tilde{g}=g$.



↑2 表示二点中插一个零值

图 9-4 重建的框图

5. 二维图象信号

为了处理图象信号, 把多分辨率推广到二维情况. 如果 $L^2(\mathbb{R}^2)$表示由二维平方可积函数构成的空间, $\{V_{2^j}\}_{j\in\mathbb{Z}}$是 $L^2(\mathbb{R})$中的一维多分辨率空间. 可以证明, $V_{2^j}^2 = V_{2^j}\times V_{2^j}$. 二维多分辨率的尺度函数为

$$\varphi(x,y) = \varphi(x)\varphi(y), \qquad (9.58)$$

其中 $\varphi(x)$、$\varphi(y)$是多分辨率空间$\{V_{2^j}\}_{j\in\mathbb{Z}}$的一维尺度函数. 与一维情况类似, 二维尺度函数的伸缩平移给出了一个 $V_{2^j}^2$空间中的正交基, 记为

$$2^{-j}\varphi(x - 2^{-j}n, y - 2^{-j}m).$$

$$\forall j \in \mathbb{Z},\ m,n \in \mathbb{Z} \qquad (9.59)$$

设 $O_{2^j}^2$是 $V_{2^j}^2$在 $V_2^{2^{j+1}}$上的正交补, 即

$$V_{2^j}^2 + O_{2^j}^2 = V_{2^{j+1}}^2, \qquad (9.60)$$

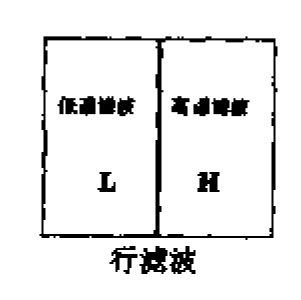
于是产生三个基本小波函数

$$\Psi^{LH}(x,y)=\varphi(x)\psi(y),$$
$$\Psi^{HL}(x,y)=\psi(x)\varphi(y),$$
$$\Psi^{HH}(x,y)=\psi(x)\psi(y). \tag{9.61}$$

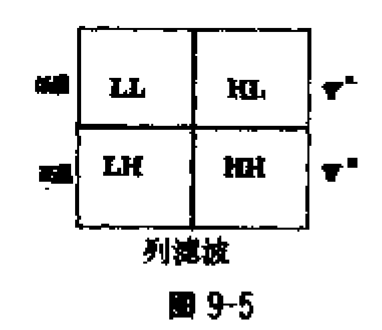
$\Omega=\{LH,HL,HH\}$表示小波函数的三个方向，由三个小波函数的伸缩平移系：

$$2^{-j}\Psi^{LH}(x-2^{-j}n,\ y-2^{-j}m),$$
$$2^{-j}\Psi^{HL}(x-2^{-j}n,\ y-2^{-j}m),$$
$$2^{-j}\Psi^{HH}(x-2^{-j}n,\ y-2^{-j}m) \tag{9.62}$$

构成$O_{2^j}^2$的正交基，并和$2^{-j}\varphi(x-2^{-j}n,\ y-2^{-j}m)$一起，构成$L^2(\mathbb{R}^2)$的正交基.



行滤波



列滤波

图 9-5

二维小波分解是和一维小波分解类似的，只是在二维情况下$D_{2^j}f$有三个，分别表示三个方向上的小波的系数，用D_{LH}^jf、D_{HL}^jf和D_{HH}^jf表示，由分辨率为2^{j+1}上的图象分解到2^j上，是先对"行"滤波，然后对"列"滤波，由於滤波器是高通和低通，因此就可得到如图 9-5 的结果，其二维小波分解示意图如图9-6所示. 在2^{j+1}上的图象$A^{j+1}f$分解，得到2^j分辨率上的四个于图象A^jf,D_{LH}^jf,D_{HL}^jf和D_{HH}^jf. 小波变换的重构过程如图9-7所示.

图9-6、图9-7只给出了一次分解和重构，其实在A^jf的分辨率上，可继续分解(称为级，或称阶). 图中$\tilde{h}$、$\tilde{g}$为对偶的低、高通滤波，对完全正交小波基来讲，$\tilde{h}=h$、$\tilde{g}=g$. 图9-8(i)和(ii)给出了一个 Lenna 头像经过一阶小波变换和三阶小波变换的结果. 在左上角上为A^jf，仍保留原图的形象，但分辨率变低，其他为不同方向的小波变换系数. 对三阶，共有10个不同尺度上的小波变换系数(包括A^jf). 小波变换如同傅里叶变换一样，其本身没有压缩效果. $2^N\times 2^N$的图象经变换后仍是$2^N\times 2^N$个系数，但此系

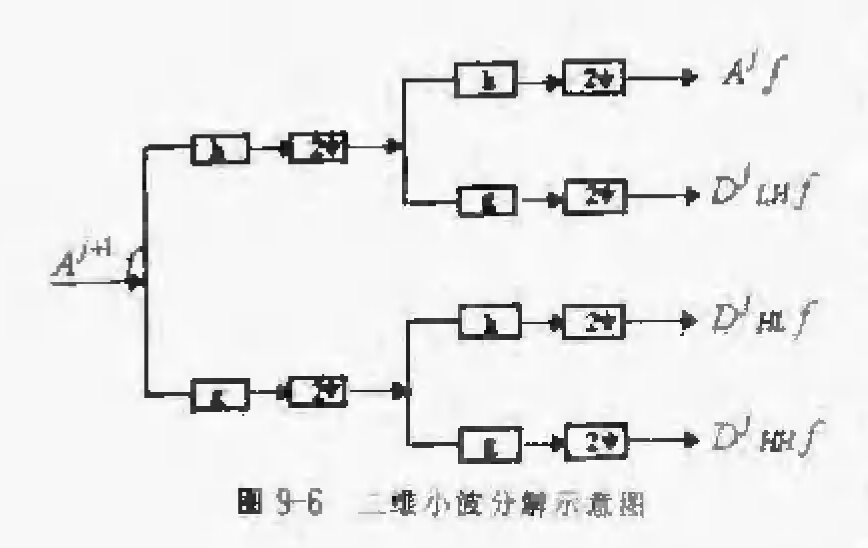


图 9-6　二维小波分解示意图

数比原始的数据有优良的性质. 如除 A^jf 之外, 其他的系数的灰阶差很小, 从而可在较小的灰阶上进行量化. 同时, 我们也可以看到, 在不同分辨率同一方向上的小波系数存在着自相似性, 即 $D^j_{HH}f$ 与 $D^{j-1}_{HH}f$、$D^j_{HL}f$ 与 $D^{j-1}_{HL}f$、$D^j_{LH}f$ 与 $D^{j-1}_{LH}f$. 能否利用这些自相似性, 在小波域中进行编码, 这就是下面要讨论的小波树方法.

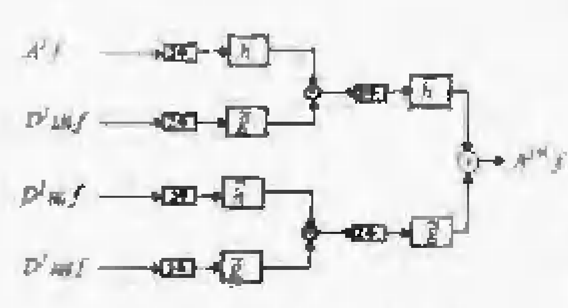


图 9-7 二维小波重构示意图



(i) 一阶小波变换



(ii) 三阶小波变换

图 9-8

§ 10　小波域的分形编码

二维小波基由三个不同方向的小波函数 $\Psi_{LH}(x, y)$、$\Psi_{HL}(x,$

$y)$、$\Psi_{HH}(x,y)$和一个尺度函数 $\Phi(x,y)$组成. 我们用下标 ω 代表三个方向 $\Omega=\{LH,HL,HH\}$中的一个. 对$2^N\times 2^N$ 的图象 I^2进行二维离散小波变换$\langle W_J,I^2\rangle$,即按框图9-6进行展开:

$$W_J=\{\Phi_{k,l}^J|0\leqslant k,\ l<2^J\}\cup\{\Psi_{\omega,k,l}^j|\omega\in\Omega,\\ J\leqslant j<N,\ 0\leqslant k,\ l<2^j\}. \tag{10.1}$$

这里 $0\leqslant J<N$,并且

$$\Phi_{k,l}^J(x,y)=2^{-J}\Phi(x-2^{-J}k,y-2^{-J}l), \tag{10.2}$$

$$\Psi_{\omega,k,l}^j(x,y)=2^j\Psi_\omega(x-2^{-j}k,y-2^{-j}l). \tag{10.3}$$

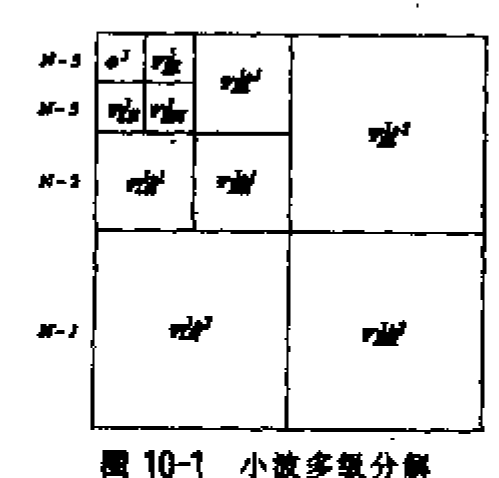


图 10-1 小波多级分解各尺度系数示意图

在二维的图象空间中,展开系数为$\langle\Phi_{k,l}^J,I^2\rangle$和$\langle\Psi_{\omega,k,l}^j,I^2\rangle$. j 为展开的级. 如果 J 的值已确定,则 $\Psi_{\omega,k,l}^j$是把原图象投影到 $V_{2^J}^2$ 分辨率和 $O_{2^J}^2,O_{2^{J+1}}^2,O_{2^{J+2}}^2,\cdots,O_{2^{N-1}}^2$上,则当 $J\leqslant j<N$,形成了多级的金字塔分解级,图10-1表示 $J=N-3$的系数分布,它们都是用图 9-6 的滤波方法,通过多级滤波得到;若 $\tilde{\Phi}_{k,l}^J$和 $\tilde{\Psi}_{\omega,k,l}^j$表示对偶尺度函数和对偶小波函数,则有

$$I^2=\langle\tilde{W}_J,\langle W_J,I^2\rangle\rangle=\sum_{k,l}\langle\Phi_{k,l}^J,I^2\rangle\tilde{\Phi}_{kl}^J\\ +\sum_{j=J}^{N-1}\sum_{\omega\in\Omega}\sum_{k,l}\langle\Psi_{\omega,k,l}^j,I^2\rangle\tilde{\Psi}_{\omega,k,l}^j. \tag{10.4}$$

有关小波域的分形编码,可参见[44]、[45].

§10.1 小波树的概念

把图10-1与图9-8(ii)作一比较. Φ^J 的系数就是分辨率比原图象小三级的系数. 它是一个很小的图象如Lenna 的头像,犹如将原图象拉到很远的距离后观察到的图象,而 Ψ_{HL}^J、Ψ_{HL}^{J+1}、Ψ_{HL}^{J+2}在图9-8

(ii)中可以看出都是在水平方向用高通滤波，而垂直方向用低通滤波得到的，它反映了垂直方向的边缘信息，从不同的分辨率上的Lenna头像本身有其自相似性，因此其小波变换后保留了原图象的空间局部特性，从而在 Ψ_{HL}^{J}、Ψ_{HL}^{J-1}和 Ψ_{HL}^{J+2}的小波系数也有其自相似性. 从图9-8(ii)中这一个分支可以明显观察到，对于其他分支 Ψ_{HH}^{J}，Ψ_{HH}^{J+1}，Ψ_{HH}^{J+2}，和 Ψ_{LH}^{J}，Ψ_{LH}^{J+1}，Ψ_{LH}^{J+2}，也具有类似的情况，只是图9-8(ii)中不太明显. 小波树是考虑 $V_{2^J\times 2^J}$分辨率上的某一个块所对应的有关小波系数上细节的组合. 假如在 I^2空间上取一个块为 $B_{K,L}^{R}I^2$，则它的信息分布在小波域中的小波树上. 为了简单起见，我们不妨用Haar小波为例. Haar小波的系数如式(9.31)所给出的. 图象在尺度函数 Φ 上的系数是四个象素点的均值，而 Ψ 的小波系数是水平方向或垂直方向的差值. 那么取块 $B_{K,L}^{R}I^2$的信息在Haar小波变换域中尺度函数 $\Phi_{k,l}^{J}$的系数上表现为左上角上位于(k,l)上的一个块，对应于 I^2中 K,L 位置的$2^{N-J}\times 2^{N-J}$块的平均灰阶，$2^{N-J}k=K$，$2^{N-J}l=L$. 如图10-2所示. 这个子块对应的小波系数又分为三个分支，合起来称为**小波树**. 小波树的根是 $\Psi_{\omega,k,l}^{J}(\omega\in\Omega)$的展开系数. 这些系数反映了块的较精细的信息(比起 Φ_{kl}^{J}的系数). 树中每个小波系数$\langle\Psi_{\omega,k,l}^{j},I^2\rangle$(其中 $j\geqslant J$)具有四个相同空间位置和方向的"孩子"，$\langle\Psi_{\omega,2k,2l}^{j+1},I^2\rangle$，$\langle\Psi_{\omega,2k,2l+1}^{j+1},I^2\rangle$，$\langle\Psi_{\omega,2k+1,2l}^{j+1},I^2\rangle$，$\langle\Psi_{\omega,2k+1,2l+1}^{j+1},I^2\rangle$. 这四个"孩于"反映了块的更精细信息. 为了方便起见，我们称小波系数$\langle\Psi_{\omega,k,l}^{j},I^2\rangle$是分辨率在 $V_{2^j}\times V_{2^j}$上的小波系数，小波树就是由三个方向上的根及其所有后代组成. 尺度函数 $\Phi_{k,l}^{J}$的展开系数$\langle\Phi_{k,l}^{J},I^2\rangle$和根系数$\langle\Psi_{\omega,k,l}^{J},I^2\rangle$有着相同的空间位置，小波树的各成员之间视觉上存在相似性，尤其是不同尺度(分辨率)同一方向上的成员体现了相似的边缘和纹理特性，我们将以树为单位，代替在空间域中的于块和父块的匹配，因而分形编码的方法延伸到小波域中进行.

§10.2 小波域中的分形编码

如同§6中"取块"操作符 $B_{K,L}^{J}$的定义一样，我们首先定义"取

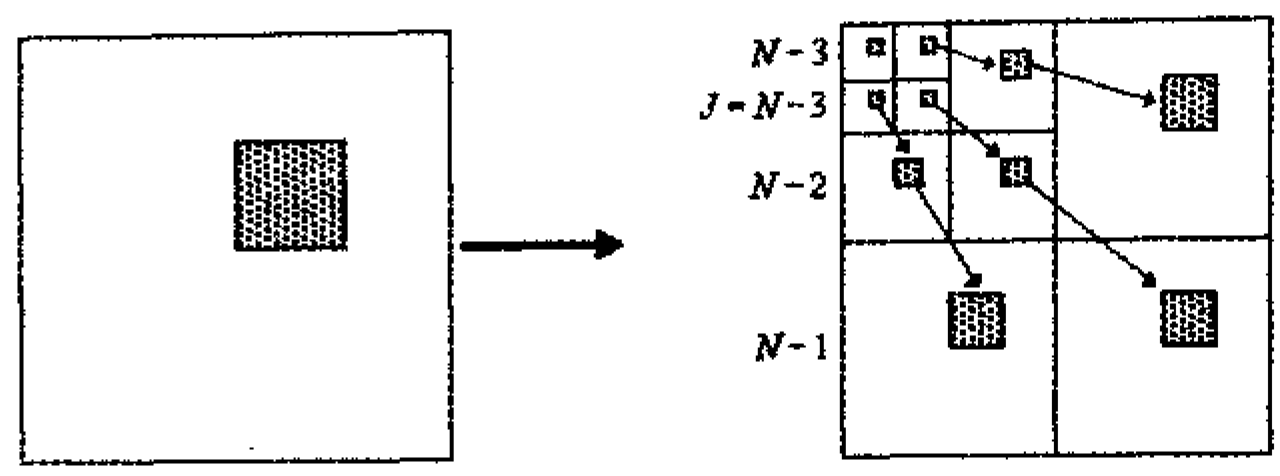


图 10-2 小波树的构造

树"操作符 $S^J_{K,L}$. 这里 J 表示分辨率，且$0\leqslant J\leqslant N$，(K,L)表示根系数在小波变换域所在的分辨率中的相对位置，且有 $0\leqslant K,L<2^J$. $S^J_{K,L}I^2_w$ 表示从图象 I^2的小波变换域 I^2_w 中，取出根在 $\Psi^J_{w,k,l}$区域的小波树. $S^J_{K,L}$的逆，为"放树"操作符$(S^J_{K,L})^*$，$(S^J_{K,L})^*$是把小波树插入全 0 的图象中，且其根在 $\Psi^J_{w,k,L}$系数区域中. 从 § 6 中知我们对$2^N\times 2^N$ 尺寸的图象 I^2，取出一个 $2^R\times 2^R$ 的子块的操作为 $B^R_{K,L}I^2$，取出$2^D\times 2^D$ 的父块的操作为 $B^D_{\pi(K,L)}I^2$，对父块的"A"操作符为邻近四个象素点灰度求平均，$L_{p(K,L)}$操作符为块的反射和旋转. 当 I^2经过小波变换后，形成了小波域，在空间域中的取块操作 $B^R_{K,L}I^2$变成了小波域中取树操作 $S^{N-R}_{K,L}I^2_w$. 例如：在 I^2中的"子块"，R 的尺寸为 $2^2\times 2^2$，则在小波域中，小波树的根只能在 $N-2$上，因为小波域中 $V_{2^{N-2}}$分辨率上，$2^2\times 2^2$的子块只剩下一个象素点，因此 $B^R_{K,L}I^2$的操作在小波域中为 $S^{N-R}_{K,L}I^2_w$，加上尺度函数的系数$\langle\Phi^{N-R}_{K,L},I^2\rangle$，在空间域上的 A 操作，表示了四点的平均，但在小波域中采用减少一个分辨率的办法进行. 如采一个块缩小了$\frac{1}{4}$，为 $V_{2^j}\times V_{2^j}$空间的小波系数，变到了 $V_{2^{j-1}}\times V_{2^{j-1}}$空间中，把这种操作记为 $\hat{A}$，称为小波域中的"平均-抽样"操作：它把小波树的"叶子"——即最精细尺度上的小波系数变为 0，其余尺度上的系数移向下一个精细尺度. 在小波域中的反射旋转变换 $\hat{L}_{P(K,L)}$将牵涉到三个小波树的方向，从而比较复杂. 下面我们用图 10-3 来说明在小波域中父树和子树的映

射关系. 在空间域中的取块和放块，在小波域中为取树和放树；空间域中的父块、子块的仿射变换，在小波域中为父树与子树的仿射变换. 图 10-3(i) 是经过五级小波变换后的系数空间 I_w^2，它的原形是一个房屋. 定义作用在原图象子块上的线性操作为 T，作用于小波树上为 $\hat{T}$. 小波树只是不考虑到尺度函数的变换，即不考虑直流成分. 图10-3(i)中的网格点块为三个方向上的小波树，其根在第二级上. 这是一个父树，它表示了房子底部的一条水平线. 对于父树的第一个操作为取树，$S_\Gamma^{N-D}I_w^2$ 为原图象 I^2 上的尺寸为 $2^D\times2^D$ 的父块在小波域 I_w^2 的图象上取出父树的操作，"Γ"表示位置(K，L)的标志；S_Γ^{N-D} 现在是取第 2→第 4 级分辨率上的网格点块，如图10-3(i)与(ii)的中间图形表示的变换. $\dot{A}$ 是去掉第 4 级分辨率上的系数而得到的结果，图 10-3 (i)、(ii)两图中间标出的图形 $\hat{L}$ 的变换共有八种，如图10-4所示. 从中选出一种最好匹配的变换作为 $\hat{L}$. "$\times g_\Gamma$"表示把变换后的灰阶乘以一个伸缩因子，如同在 § 6 式(6.7)中 $C_{K,L}$的因子一样，为灰阶的伸缩因子. 而平移因子 $h_{K,L}$ 可以不考虑，这是因小波树中已没有平移的直流成分，直流都已在 $\Phi_{K,L}^j$中考虑了. $(S_\Gamma^{N-R})^*$ 的"放树"是在小波域中进行的，结果为图 10-3(ii) 所示，它与(ii)中的深网点块组成的子树相匹配.

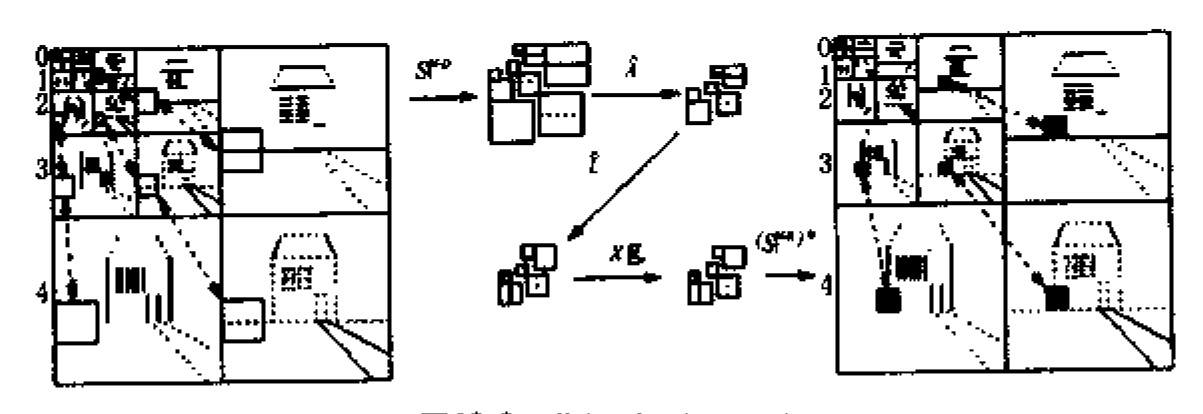


图10-3 父树到子树的映射

图 10-4 表示 8 种变换 $\hat{L}_p$，因为对角线反射和±90°旋转使水平分量和垂直分量恰好对换一下，所以需要把 Ψ_{LH}和 Ψ_{HL} 方向上的小波系数也对换一下. 因此，只有当小波基是对称或反对称的情况下，8 种对称操作才有意义. 此时，子块的中心方能保持不变. 因

此，在§ 9 介绍的小波种类中只有 Haar 小波、频率域是矩形小波、Meyer 小波、样条小波和双正交小波才能选作分形编码所需要的小波变换. 又因频率域是矩形小波、Meyer 小波和样条小波在时间域中无界，因此常用的小波基为 Haar，或 Adelson[42]，9点双正交小波和 Antonini[43]的7/9小波基.

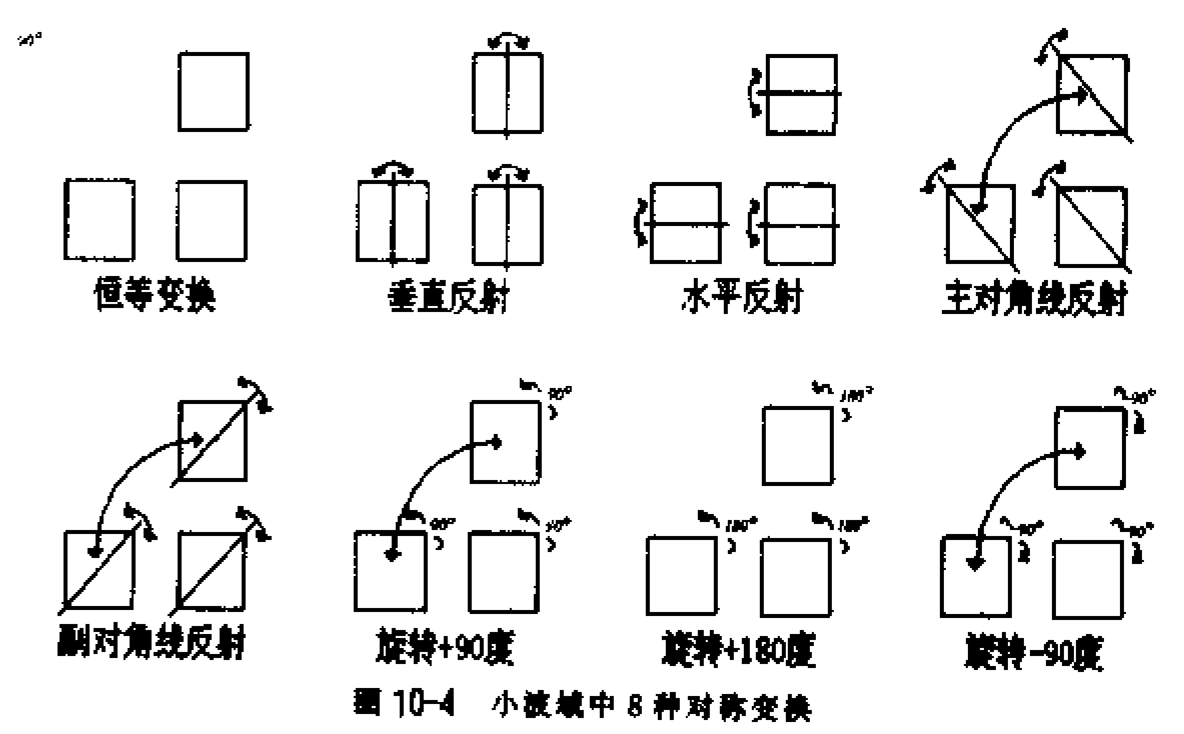


图 10-4 小波域中 8 种对称变换

与传统的空间域分形编码相似，小波域中是用小波父树 $S_\Gamma^{N-D}I_\omega^2$ 来逼近 $S_\Gamma^{N-R}I_\omega^2$. 类似于(6.7)式，只是 $h_{K,L}=0$，不考虑直流分量，有

$$S_\Gamma^{N-R}I_\omega^2 \approx g_\Gamma \hat{L}_{P(\Gamma)}\hat{A}^{D-R}S_{\pi(\Gamma)}^{N-D}I_\omega^2. \tag{10.5}$$

对于每一棵小波子树，寻找最佳的小波父树，使(10.5)式左右两边尽可能接近，保存最佳的父树位置 π、旋转反射类型 P、伸缩因子 g_Γ. 与传统的不同的是不需要保存偏移量 h. 由于映射是在整个小波树上进行，搜索的范围小，子树和父树相对的面积比较大. 但对于尺度小于 $N-R$ 的小波系数，必须要保留.

解码方法也和传统的空间域中相同. 精细的子树 $S_\Gamma^{N-R}I_\omega^2$ 是由粗糙的父树 $S_\Gamma^{N-D}I_\omega^2$ 映射得到的. 为了方便起见，设 $D=R+1$. 我们从最粗糙的 J 尺度系数开始，通过映射产生下一精细的 $J+1$ 尺度的系数，依次类推，直到产生最精细的 $N-1$ 尺度的系数. 以图 10-

3 为例，我们事先已知尺度 0,1,2 中的系数，先以父树 $S_{\Gamma}{}^{2}I_{\omega}^{2}$ 的根构成码书，得到子树 $S_{\Gamma}{}^{3}I_{\omega}^{2}$ 的根，从(10.5)式可得

$$S_{\Gamma}^{J}I_{\omega}^{2}=g_{\Gamma}\hat{L}_{P(\Gamma)}\hat{A}S_{\pi(\Gamma)}^{J-1}I_{\omega}^{2}, \tag{10.6}$$

这样得到尺度 3 上的所有小波系数. 再次运行(10.6)式，得到下一精细尺度的所有小波系数. 经过 $N-J$ 次运用，我们得到了全部小波系数.

定理 1 设 I^2 是 $2^N\times 2^N$ 象素的图象. 假定已知尺度小于 J 的所有小波系数，且根位于尺度 J 的小波树满足 $S_{\Gamma}{}^{J+1}I_{\omega}^{2}=g_{\Gamma}\hat{L}_{\Gamma}\hat{A}S_{\Gamma}^{J}I_{\omega}^{2}$，则通过 $N-J$ 次迭代，可以得到 I^2 的全部小波系数.

证明略.

上述定理表明，小波域的子树分形解码过程必将在有限步内完成. 迭代的过程是无条件收敛的，不需要强制要求 $|g_{\Gamma}|<1$.

§10.3 小波域的分形编码方法的实现

1. 子树的尺寸固定情况下的实现

把小波域图象分割成大小相同的子树 $S_{K,L}^{N-R}I_{\omega}^{2}$，其中 $(K,L)\in R_I$，$R_I=\{(2^R m,2^R n)\,|\,0\leqslant m,n<2^{N-R}\}$. 父树 $S_{K,L}^{N-R}I_{\omega}^{2}$ 的选取与空间域父块的选取略有差别. 如果把树根之间的距离定义为树的间距，相距一个象素点距离的父树集对应到空间域，变成相距一个父块宽度 2^D 的父块集. 但在空间域中相距小于一个父块宽度的父块集，在小波域中确无法表示. 所以父树的选择范围比在空间域中的父块少得多. 只有相距 $K2^D$ 象素 $(K\in\mathbb{Z}^+)$ 的父块集，才有对应的父树集.

经过"平均-抽样"的算子 $\hat{A}$ 和"反射旋转"算子 $\hat{L}$ 变换后，父树与子树近似程度仍用 §6 的式(6.9)、(6.10)、(6.11)进行计算，得到最优的父树. 由于现在小波变换和分形两者同时存在，我们用一维的图来表示两者是如何共存的. 图 10-5 是一维小波域的分形量化示意图，它是通过五级小波变换得到 φ^0、ψ^0、ψ^1、ψ^2、ψ^3、ψ^4 的小波系数. 图中的深色块表示是小波系数，白色块表明需要进行分形

编码的部分，图中加了粗线框的"子树"$S_1^{-2}f$，是由"父树"$S_1^{-1}f$ 映射得到的．最后保存的是经过标量量化的小波系数和分形编码得到的映射系数，包括父树位置、反射旋转变换的类型和伸缩因子．

解码时，先恢复标量量化编码的小波系数，这是粗糙尺度的系数；然后利用迭代公式(10.5)，逐次恢复 $N-R$、$N-R-1$、……、$N-1$尺度的小波系数，经$(R-1)$步就可恢复全部小波系数．

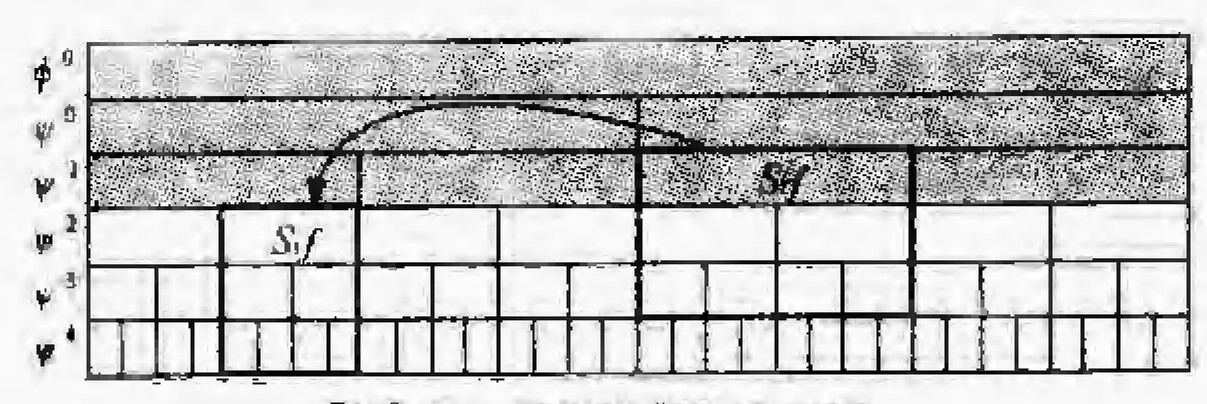


图 10-5　一维小波域分形量化示意图

2. 四叉树分割的分形编码

小波域上的分形编码也可以采用四叉树分割的方法．只是有一点需要注意，若根位于尺度 $N-R$ 的子树 $S_{K,L}^{N-R}I_\omega^2$ 在根的尺度上只有三个方向上的三个系数构成，假设为 $C_{HL,K,L}^{N-R}$、$C_{LH,K,L}^{N-R}$、$C_{HH,K,L}^{N-R}$ 构成，如果在子树与父树进行匹配搜索中找不到相匹配块，即因匹配误差超过阈值而对子树进行分割时，$C_{HL,K,L}^{N-R}$、$C_{LH,K,L}^{N-R}$和 $C_{HH,K,L}^{N-R}$再也不能分割下去，从而只能把子树根的分辨率提高一步进行，把原子树 $S_{K,L}^{N-R}I_\omega^2$ 变为四个"孩子"：$S_{2K,2L}^{N-(R-1)}I_\omega^2$、$S_{2K+1,2L}^{N-(R-1)}I_\omega^2$、$S_{2K,2L+1}^{N-(R-1)}I_\omega^2$ 和 $S_{2K+1,2L+1}^{N-(R-1)}I_\omega^2$，并留下三个系数 $C_{HL,K,L}^{N-R}$、$C_{LH,K,L}^{N-R}$和 $C_{HH,K,L}^{N-R}$不能分割，新的子树可以继续分割，直到匹配误差小于阈值或者子树大小达到规定的尺寸为止．因此，小波系数的范围不像图 10-5 那样与分形编码分得十分清楚．还是以一维为例，图 10-6 表示小波系数与分形编码的小波树是如何"共处"的．图中的深色块表明保留的小波系数，加了粗线框的白色块表明参与小波树编码，保存的是量化后的映射系数，并非原块中的小波系数．图 10-5 和图 10-6 的主要区别是后者的子树大小可以自动调整，因此前者进行量化编码时

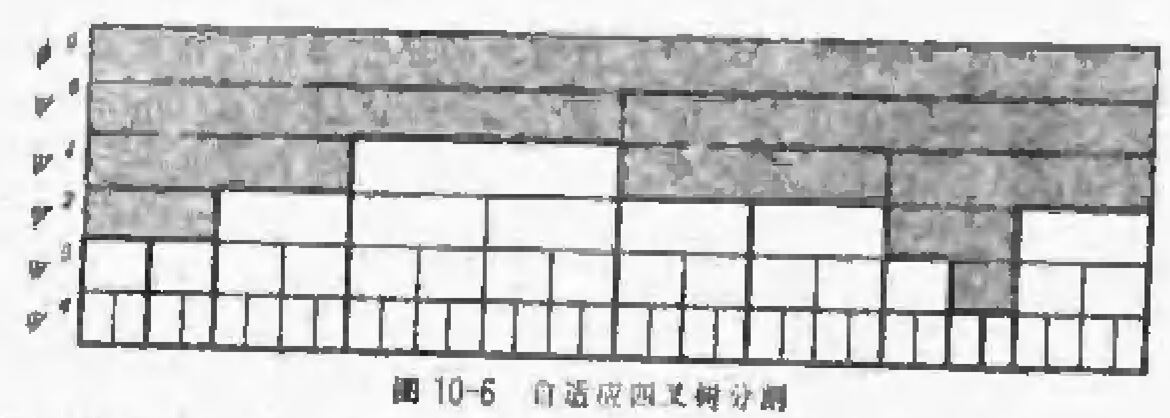


图 10-6 自适应四叉树分割

只是尺度小于 $N-R$ 的全部小波系数，而后者却可能有参差.

分形解码时，仍然先恢复编码后的小波系数，然后利用迭代公式(10.5)，逐次恢复 $N-R_M$、$N-R_M+1$、……、$N-1$ 尺度中参与小波树分形编码的小波系数. R_M 是最大的子树根所在的位置.

在四叉树分割方法中，动态地调整子树大小，其中阈值的选取对压缩效果影响很大，阈值 ε 越大，子块越容易得到匹配，剩下的小波系数也越少(遗留的根系数少)，编码效率很高，但此时误差也较大；相反，阈值越少，子块越不容易满足匹配条件，于是子块越分越多，所需的分形系数也多，同时剩下的小波系数也越多，这时，编码效率就低，但误差小，质量较高. 在实际应用中，可根据总比特率的要求来调整. 如总码流为 B 比特，分形系数为 B_f 比特表示，余下的 $B_w=B-B_f$ 表示小波系数. 这样也可根据 B 的要求对小波系数进行量化或对阈值 ε 进行调节.

总的算法的步骤为：设压缩后图象的总码率为 B，首先，对图象进行小波变换，小波基可以采用对称或反对称的现有小波基函数；第二，对小波变换后的图象进行分形编码；第三，令阈值为 ε，用四叉树分割的方法，对尺度 $\geqslant N-R$ 的小波系数进行分形编码，设保存分形系数所需的比特数为 B_f；第四，对余下的小波系数进行量化编码，余下的小波系数中，除了尺度 $<N-R$ 的全部小波系数，还有因四叉树分割而保留下来的部分尺度 $\geqslant N-R$ 的系数，用剩下的比特数 $B_w=B-B_f$ 对这些系数进行分层量化编码；第五，计算如此编码造成的均方误差，用二分法搜索最佳阈值.







(f)Haar 基　PSNR=26.1db　　(g)Antonini 基　PSNR=27.5db

图 10-7　自适应四叉树分解(压缩100倍)

解码的步骤如下：

1. 先恢复分层量化编码的小波系数

2. 通过分形解码，恢复其余小波系数。利用已知的小波系数和分形系数，运用(10.5)式迭代后，恢复所有小波系数。

3. 小波反变换。

利用小波域的分形编码，可得到高质量、高压缩比的结果。图10-7为512×512的灰度图象，其压缩比可达100倍。

§11　分形与离散余弦变换(DCT)结合的方法

离散余弦变换是N. Ahmed 等人在1974年提出的正交变换方法。它常被认为是对语音和图象信号进行变换的最佳方法。为了工程上实现的需要，国内外许多学者花费了很大精力去寻找或改进离散余弦变换的快速算法。由于近年来数字信号处理芯片(DSP)的发展，加上专用集成电路设计上的优势，这就牢固地确立了离散余弦变换(DCT)在目前图象编码中的重要地位，成为H.261、JPEG、MPEG等国际上公用的编码标准的重要环节。以DCT方法

本身在高压缩比时，其方块效应十分明显，因而在有些高存储量的图象信息压缩、低码率的图象通信中，DCT 方法暴露出质量不高的缺点，分形方法虽然具有高的压缩比，但计算量大，因而不能达到实用. DCT 与分形相结合的方法是企图取各自的长处，来达到高质量和高压缩比的目的.

§11.1 余弦变换的定义

以$\{x(m)\}$表示 M 个其值有限的一维实数信号序列的集合，$m=0,1,\cdots,M-1$. 它的一维 DCT 定义为

$$Y(k)=\sqrt{\frac{2}{M}}c(k)\sum_{m=0}^{M-1}x(m)\cos\frac{2m+1}{2M}k\pi.$$

$$(k=0,1,\cdots,M-1)\quad(11.1)$$

一维的余弦反变换(IDCT)定义为

$$x(m)=\sqrt{\frac{2}{M}}\,c(m)\sum_{k=0}^{M-1}Y(k)\cos\frac{2m+1}{2M}k\pi,\qquad(11.2)$$

其中

$$c(k),c(m)=\begin{cases}1/\sqrt{2}, & \text{当 } k,m=0 \text{ 时};\\ 1, & \text{当 } k,m=1,2,\cdots,M-1 \text{ 时}.\end{cases}\qquad(11.3)$$

(11.1)可以写成矩阵形式：

$$\boldsymbol{Y}=\boldsymbol{A}\boldsymbol{X}.\qquad(11.4)$$

其中 $\boldsymbol{A}$ 为 DCT 变换矩阵：

$$\boldsymbol{A}=\left\{\sqrt{\frac{2}{M}}\left\{c(k)\cos\frac{(2m+1)k\pi}{2M}\right\}\right\}_{M\times M}.$$

$$(k,m=0,1,\cdots,M-1)\quad(11.5)$$

可以验证，DCT 矩阵 $\boldsymbol{A}$ 是一个正交矩阵，但不是一个对称矩阵. IDCT 的矩阵为

$$A^{-1} = A^T = \left\{ \sqrt{\frac{2}{M}} \left\{ c(k)\cos\frac{(2m+1)k\pi}{2M} \right\} \right\}_{M\times M}.$$

$$(m,k = 0,1,\cdots,M-1) \quad (11.6)$$

它除了行和列的序号互换外，与(11.5)式完全相同. 也可用矩阵形式表示：

$$X = A^T Y. \quad (11.7)$$

数字图象$\{x(m,n)\}$是具有 M 行、N 列的一个矩阵. 为了同时减弱或去除空间域的相关性，将图象从空间域变换到频率域. 二维的 DCT 表达式为

$$Y(k,l) = \frac{2}{\sqrt{MN}} c(k)c(l) \cdot \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} x(m,n)\cos\frac{(2m+1)k\pi}{2M}\cos\frac{(2n+1)l\pi}{2N}, \quad (11.8)$$

其中 $k=0,1,\cdots,M-1$；$l=0,1,\cdots,N-1$. 而它的反变换为

$$x(m,n) = \frac{2}{\sqrt{MN}} \sum_{k=0}^{M-1}\sum_{l=0}^{N-1} c(k)c(l)Y(k,l) \cdot \cos\frac{(2m+1)k\pi}{2M}\cos\frac{(2n+1)l\pi}{2N}, \quad (11.9)$$

其中 $m=0,1,\cdots,M-1$；$n=0,1,\cdots,N-1$.

式(11.8)、(11.9)都可以拆成两个一维的 DCT，如先对“行”进行一维 DCT，再对“列”进行一维 DCT，当然也可用并行的 DCT.

在现有的图象压缩标准中，都是把图象划分为很多 8×8 的子块，对每个 8×8 的子块进行 DCT 变换. 设 $x(m,n)$是任一个 8×8 的子块，则子块的 DCT 为

$$Y(k,l) = \frac{1}{4} c(k)c(l) \sum_{m=0}^{7}\sum_{n=0}^{7} x(m,n) \cdot \cos\frac{(2m+1)k\pi}{16}\cos\frac{(2n+1)l\pi}{16}, \quad (11.10)$$

其中 $k=0,1,\cdots,7$；$l=0,1,\cdots,7$，

$$c(k),c(l)=\begin{cases}1/\sqrt{2}, & \text{当 } k,l=0 \text{ 时;}\\ 1, & \text{当 } k,l \text{ 为其他.}\end{cases}$$

经 DCT 变换后，每个子块成了 64 个 DCT 的系数 $Y(k,l)$. 对 $Y(k,l)$进行量化处理，设 $Q(k,l)$是量化步长，根据视觉的效应，$Q(k,l)$对不同的频率(k,l)的量化步长是不同的，量化结果为

$$Y_Q(k,l)=\text{round}\left(\frac{Y(k,l)}{Q(k,l)}\right). \tag{11.11}$$

这里，round 函数是四舍五入取整函数. 量化后的直流系数，$Y_Q(0,0)$和 63 个交流系数分别处理，直流系数与前一个子块的直流系数进行预测编码，保存其差. 量化后的交流系数被转换成"Z"形序列，用熵编码化为码流处理. 这就是目前标准的图象压缩中所用的方法. 注意在量化步长 $Q(k,l)$中有一个调节因子 Q_P，改变 Q_P 可以同时改变 $Q(k,l)$的步长，对一切 k,l 成立，因而 Q_P 的大小可控制压缩比和质量. 在固定码流时，Q_P 成为 DCT 编码的控制因子. 其编码和解码的框图已在 §5中图 5-2讨论了，这里不再重复.

§11.2 DCT 变换系数的分形编码方法

如果把原始图象划成两种类型的块，一种为 8×8 象素组成的子块，另一种是 16×16 的父块，它们都被分别进行 DCT 变换，用 $F_R(k,l)$和 $F_D(k,l)$表示. 而 $F_R(k,l)$根据其交流系数分为两种类型，一种为简单子块，一种为复杂子块. 假设 T_N 为给定的阈值，则有

$$|F_R(0,1)|+|F_R(1,0)|+|F_R(1,1)| =\begin{cases}<T_n, & \text{当 } F_R \text{ 为简单子块;}\\ \geq T_n, & \text{当 } F_R \text{ 为复杂子块.}\end{cases} \tag{11.12}$$

$F_R(0,0)$是直流成分，而 $F_R(0,1)$、$F_R(1,0)$、$F_R(1,1)$是比较低频率的交流成分. 在一般的 DCT 变换块中，频率的分布形状比较相似，低频成分的幅度较大，而高频成分幅度较小，因此，如果比较低的频率的交流成分的幅度小，说明这个子块几乎是一个平坦区，此

时我们只要保留 $F(0,0)$ 的成分，就保留了其主要信息量，对此子块的处理比较方便，只需存几个比特信息. 对子复杂子块，可采用分形自仿射变换

$$F_R^{K,L} \approx C_{K,L} \cdot L_{P(K,L)} \cdot A \cdot F_D^{\pi(K,L)} + h_{K,L} \cdot I. \quad (11.13)$$

这里 $L_{P(K,L)}$ 为反射旋转因子，A 为平均抽样因子，$C_{K,L}$ 为伸缩因子，$h_{K,L}$ 为平移因子，I 为全 1 的 8×8 矩阵，大写的字母 K,L 为子块左上角的位置. $\pi(K,L)$ 为父块的位置，其自仿射变换的系数求法与 §6 中相同，只是现在 F_R、F_D 都是在 DCT 变换系数中进行. 设 $\tau(F_D(k,l))$ 代表(11.13)式左面的变换，那么其自仿射的误差为

$$E = \sum_{l=0}^{7}\sum_{k=0}^{7} |F_R(k,l) - \tau(F_D(k,l))|^2. \quad (11.14)$$

如果 $E<\varepsilon$，则可保留 $C_{K,L}$、$h_{K,L}$、$L_{P(K,L)}$ 和 $\pi(k,l)$ 系数. 其误差函数为

$$E(k,l) = F_R(k,l) - \tau(F_D(k,l)). \quad (11.15)$$

如果误差 $E(k,l)$ 仍比较大，则可继续对误差块 $E(k,l)$ 进行编码，得到其补偿误差值的无失真编码，以保证在分形解码时，恢复 DCT 系数的误差补偿. 如果误差 $E(k,l)$ 比较小，则(11.13)的分形编码已达到恢复 DCT 系数的目的.

在解码时，首先从得到的分形系数中进行任意初始图象的分形迭代. 类似于 §6 中的方法，只是因为父块和子块都是在 DCT 变换系数中进行，因此迭代一次要用 IDCT(逆变换)回到原始图象，由於 DCT 系数中的误差也可作为补偿参加到迭代中去，从面改进图象的质量. 系数中还有补偿的误差，也要加入到分形编码中进行迭代. 在 DCT 变换系数中的分形编码[47]的优点有：1. 在空间域中比较难找到自相似性的子块和父块，在 DCT 变换系数中就比较容易找到其自相似性，这是因为频率域的形状都是低频端振幅大，高频端幅度小；2. 经过许多学者的研究、改进，DCT 的变换的速度比较快；3. 加入了补偿项，改进了图象压缩的质量.

§11.3 分形为主的DCT补偿方法

分形编码最常用的方法是四叉树分割的方法.在实际应用中，此方法存在以下几个问题:1.编码的时间很长,尤其当对图象质量要求高时,所需时间更长,虽然可用硬件来完成,但毕竟代价很高;2.压缩质量高低与最小子块的大小很有关系,子块的尺寸取得小时质量才可能提高,但此时压缩比提不高,子块取得大时,压缩比大而质量不高;3.通过容许误差 ϵ 来调节压缩比和质量时,当容许误差小时,编码时间大大增加,但不能保证质量能提高很多,因此不容易控制图象质量和压缩比.

分形与DCT结合的思想,是类似于MPEG标准中的P帧的运动矢量匹配后的运动补偿块作DCT编码[48].它把运动矢量搜索看作是同一幅图象中父块到子块的匹配.即对图象先用四叉树的分形编码方法,若图象的质量达不到要求,则对经过分形编码后还没有达到要求的误差部分,再用DCT进行编码和补偿,而DCT的图象质量和压缩比控制是利用量化因子 Q_P,因而对 ϵ 和 Q_P 两个量的控制使图象的质量和压缩比得到较好的调节.而在解码时把误差作为迭代函数系统中的偏移量处理,具体做法建立在四叉树编码基础上.

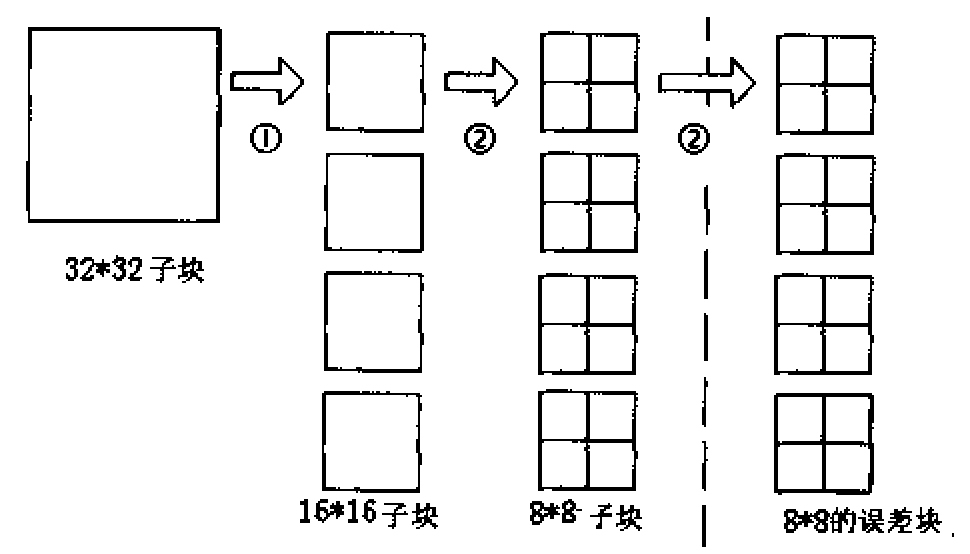


图11-1 四叉树分割和误差块

假设图象I^2被分为R_{max}的子块，每个R_{max}的尺寸为$2^{R_{max}\times R_{max}}$，$R_{max}$是可考虑的最大尺寸的子块，与它相匹配的父块$D_{max}$有尺寸$2^{D_{max}}\times 2^{D_{max}}$. 按(6.7)式，要寻找一个与位置在$K,L$上的与子块相匹配的父块和参数$C_{K,L}$、$h_{K,L}$，满足

$$B_{K,L}^{R\max}I^2 \approx C_{K,L}L_{P(K,L)}A^{D\max - R\max}B_{\pi(K,L)}^{R\max}I^2 + h_{K,L}B_{K,L}^{R\max}\cdot \boldsymbol{I}. \tag{11.16}$$

定义(11.16)式的右边为$W_{K,L}(B_{\pi(K,L)}^{D\max}I^2)$. 如果

$$E_{\min} = \|B_{K,L}^{R\max}I^2 - W_{K,L}(B_{\pi(K,L)}^{D\max}I^2)\|^2 < \varepsilon, \tag{11.17}$$

ε为预定的阈值，则记下$C_{K,L}$、$h_{K,L}$、$L_{P(K,L)}$和$\pi(K,L)$，这时在K,L位置上的R_{max}编码就完成了. 此时的编码效率最高，压缩比也最大. 但是，不是所有的最大尺寸的子块都能找到它相匹配的父块，因此，对于不满足(11.17)式的子块，必须通过四分法得到$R_1=\frac{R_{max}}{4}$，$D_1=\frac{D_{max}}{4}$，仍按(11.16)式搜索，即

$$B_{K,L}^{R_1}I^2 \approx C_{K,L}L_{P(K,L)}A^{D_1-R_1}B_{\pi(K,L)}^{D_1}I^2 + h_{K,L}\cdot B_{K,L}^{R_1}\cdot \boldsymbol{I}.$$

只要E_{min}不满足$E_{min}<\varepsilon$，则$R_2=\frac{R_1}{4}$，$D_2=\frac{D_1}{4}$，继续划分和搜索，其结果会造成§11.3开始时提到的问题1～3.

现在定义R_{min}为一个给定的最小子块. 不失一般性，可以利用现有的DCT算法中的尺寸8×8. 当四叉树分割到$R=R_{min}$时，仍不满足$E_{min}<\varepsilon$的条件，就不再分下去了，记下满足E_{min}时的参数$C_{K,L}$、$h_{K,L}$、$L_{P(K,L)}$和$\pi(K,L)$，不满足部分用一个误差块来补偿，假设

$$E_{\min} = \|B_{K,L}^{R\min}I^2 - W_{K,L}(B_{\pi(K,L)}^{D\min}I^2)\|^2 > \varepsilon.$$

令其误差块为

$$E_{K,L}^{R\min} = B_{K,L}^{R\min}I^2 - W_{K,L}(B_{\pi(K,L)}^{D\min}I^2), \tag{11.18}$$

具有R_{min}尺寸8×8. 令$F(E_{K,L}^{R\min})$为$E_{K,L}^{R\min}$的DCT变换，$F_q(E_{K,L}^{R\min})$为DCT变换系数的量化系数，那么在以分形为主的DCT变换方法中，那些能找到其相匹配父块的子块都用分形方法进行编码，那些

找不到与其匹配父块的子块，不再继续用四叉树方法划分了，而是保留其匹配误差，并用DCT进行匹配误差块的编码. 这样既防止了继续四叉树划分时编码时间很长的问题，又用DCT补偿的方法提高了图象的质量. 同时，由于分割的阈值 ϵ 和DCT的量化因子 Q_P 的调节，使图象的质量和压缩比得到了控制. 其具体划块如图11-1所示. $R_{max}=2^5\times2^5$，$R_{min}=2^3\times2^3$.

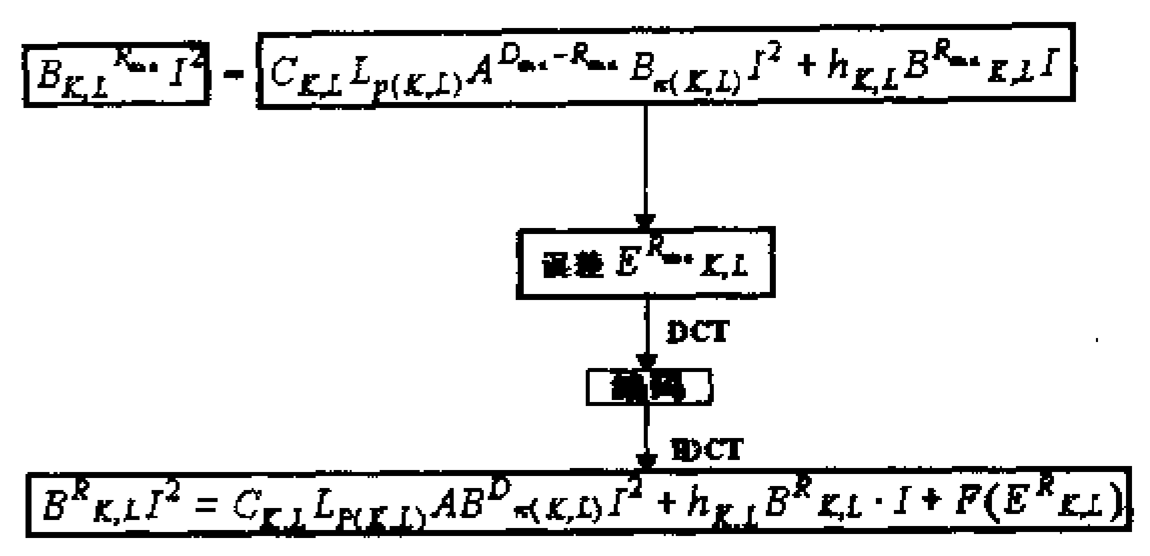


图11-2　编码解码示意图

分形与DCT结合的解码方法也是一个迭代过程. 它的步骤为：

1. 读入各误差块的DCT变换参数，并根据参数先解出DCT部分的编码，得到误差数据 $F(E_{K,L}^{R_{min}})$，为逆余弦变换的误差块.

2. 其分形迭代公式为

$$B_{K,L}^R I^2 = C_{K,L} L_{P(K,L)} A^{D-R} B_{\pi(K,L)}^D I^2 + h_{K,L} \cdot I + F(E_{K,L}^R). \tag{11.19}$$

因为 $F(E_{K,L}^R)$ 完全与伸缩因子无关，仅是一个平移因子，所以不影响迭代的收敛性. 而 $F(E_{K,L}^R)$ 也只有对那些 $E_{K,L}^R$ 误差不为0的 R_{min} 大小的子块进行，因而计算量不大. 再由於迭代是自仿射形式，它对整个图象的恢复起了很重要作用. 图11-2表示其编码和解码的示意图.

由於DCT本身具有较快的编码速度，加上分形编码本身具有较高的压缩比，可以通过调节四叉树分割阈值和DCT量化得到比

较好的综合结果.

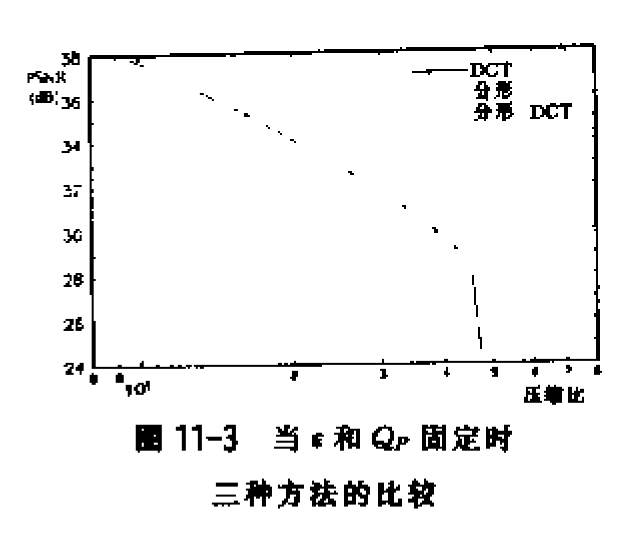


图 11-3 当 ϵ 和 Q_P 固定时三种方法的比较

图 11-3 表示三种方法在压缩比和质量上的关系. 在低压缩比时,DCT 方法最佳;但当压缩比超过35～40db 后,分形编码和"以分形为主的 DCT 补偿编码"占了优势,而"以分形为主的 DCT 补偿编码"方法又优于纯分形编码的方法.

图 11-4 表示三种方法所花费的时间."分形为主的 DCT 编码"补偿方法虽然没有单纯的 DCT 方法快,但比起单纯的分形编码方法要快得多.

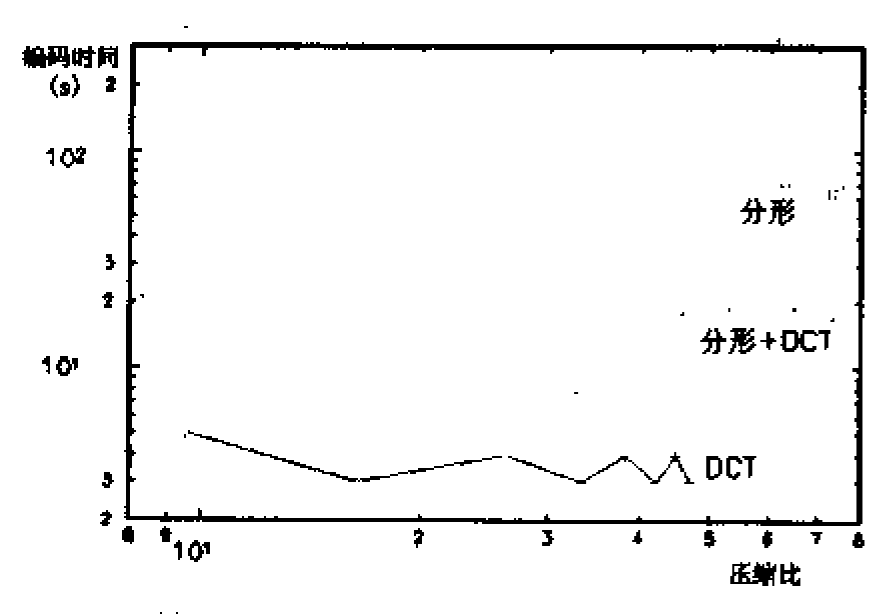


图 11-4 三种方法的编码时间和压缩比关系

其实,如果适当地调节 ϵ 和 Q_P,可以使图 11-3 达到更好的效果. 因为当 ϵ 取得很小时,会出现任何子块都不能找到与其相匹配的父块,于是所有的子块与最佳匹配父块的误差都需要用 DCT 进行补偿,此时编码的结果与 DCT 相当. 即在低压缩比时可用这种方法,使其编码质量与经典的 DCT 方法一样,只是因在初始时需要先用分形搜索,其时间略长,而且需要多保留一些分形编码系数,因而压缩比略低于 DCT 方法. 当 ϵ 取得略大,这时分形编码的

成分多了，只有少量的子块找不到匹配父块，这样压缩比可提高，而又因 DCT 对误差的补偿，使图象质量高于纯分形方法，只要找到 ϵ 和 DCT 量化因子的关系就可使图象编码时自动达到最优的状态.



图11-5 不同方法压缩图象的结果

图 11-4 所示的运行时间是在 *Sun*10工作站上运行来进行比较的. 纯分形的方法在低压缩比时反而需要更长的时间,这是因为低压缩比需要存储更多的自仿射变换系数,这意味着四叉树分割得越细. 而"分形为主的 *DCT* 补偿方法"其编码时间与压缩比无关.

图 11-5(i)～(vi)表示用不同的方法进行压缩的质量和时间. *LIFSM* 表示分块的分形方法,*LIFSM*+*DCT* 表示"以分形为主的 *DCT* 补偿方法".

参考文献

[1] 曼德勃罗特,《大自然的分形几何学》(中译本)的序(1997.1.9),陈守吉、凌复华译,上海远东出版社,1998.

[2] B. B. Mandelbrot, *Les objets fractals: forme, hasard et dimension*, Flammarion,1975.

[3] B. B. Mandelbrot, *Fractals: form, chance, and dimension*, Freeman, 1977.

[4] Y. Fisher. (ed.) *Fractal Image Compression: Theory and Application*, Springer, 1995.

[5] K. Falconer, *Fractal Geometry*, Wiley & sons,1990.

[6] B.B. Mandelbrot, *The Fractal Geometry of Nature*, freeman,1983

[7] J. Huchinson, *Fractals and Self Similarity*, *indiana Univ.* Math. J. 30.1981.

[8] M. F. Barnsley, et. al, *Iterated function systems and the global construction of fractals*, Proc. Roy. Soc. London, Ser. A 399, 1985.

[9] M. F. Barnsley, *Fractals Everywhere* (*2 ed.*), Academic Press, 1993.

[10] J. Elton, *An ergotic theorem for iterated maps*, Egotic Theory Dynamical System, 7, 1987.

[11] M. F. Barnsley, et. al, *The Desktop fractal Design System Handbook*, Academic Press, 1992.

[12] M. F. Barnsley, et. al, *Fractal image compression*, AK Peter, 1993.

[13] Oden, *Applied Functional Analysis*, 1975.

[14] 顾其均,图象分形编码与压缩技术,《分形几何讲义》,非线性科学研讨班讲义,1994.

[15] A. P. Pentland, *Fractal-based description of nature scenes*, IEEE Trans. MAMI-6(6),1984.

[16] 齐东旭,《分形及其计算机生成》,科学出版社,1994.

[17] M. F. Barnsley, *Fractals Functions and Interpolation*, Constractive Approximation, Vol. 2, 1986.

[18] 吴平南,《数据压缩的原理与应用》,电子工业出版社, 1995.

[19] M. F. Barnsley, A. D. Sloan, *A Better Way to Compress Image Byte*, Jan. 1988.

[20] A. E. Jacquin, *A Novel fractal Block-Coding Technique for Digital Image*, *Proceedings of ICASSP IEEE International conference on ASSP*, Vol. 4 1990.

[21] A. E. Jacquin, *Fractal Image Cnding: A Review*, *Proceedivg of the IEEE* Vol. 81 No. 10 1993.

[22] Y. Fisher, *Fractal Image Compression*, Fractals Vol. 2 No. 3 1994.

[23] Y. Fisher, *Fractal Encnding: Theoy and Application to Digifal Image*, Springer Verlag, New York 1994.

[24] E. W. Jacobs, Y. Fisher, R. D. Boss, *Image Compression: A Study of the Itrated Transform Method*, Signal Processing Vol. 29, 1992.

[25] T. Bedford, F. M. Dekking, M. S. Keane, *Fractal Image Coding Techniques and Contraction operators*, Nieuw Arch. Wisk (4) 10, 3, 1992.

[26] A. E. Jacquin, *Image Cnding Based on a Fractal Theoy of Iterated Contractive Markov Operators*, *Part I、II: Theoretical Foundation*, Report Math. 91389—016 Georgia Institute of Technology, 1989.

[27] A. e. Jacquin, *Fractal Image Coding Based on A Theoy of Iterated Contractive Image Transformafions*, *IEEE Trans*. Image Proessing Vol. 1 1992.

[28] H. Lin, A. N. Venetsanopoulos. *A Pyramid algorithm for Fast Fractal Image Compression*. Proceedings of ICIP. 95 IEEE International Conference on Image Processing 1995.

[29] D. Saupe, R. A. Hamzaoui, *A Guided Tour of The Fractal Image Compression Literature*, *New Directions for Fractal Modeling in Computer Graphics*, *J. Hart*, *ACM Siggraph Course Notes* 13, 1994.

[30] F. M. Dekking, *An Inequality for Pairs of Martingales and Its Applications to Fractal Image Coding*, *Technical Report* 95—10 *Faculfy of*

Technical Mathematics and Informatics Delft university of Technology, 1995.

[31] F. M. Dekking, Fractal Image Coding: *Some Mathematical Remarks on Its Limits and Its Prospects, Thechnical Report, Faculty of Technical Mathematics and Informatics Delft University of Technology*, 1995.

[32] H. Lin, A. N. Venetsanopoulos, *Incorporating Nonlinear Contractive functions into Fractal Coding, Proceedings of International Workshop on Intelligent Signal Procssing and Communication System*, 1994.

[33] Y. Fisher, S. Menlove, *Fractal Encoding With HV Partitions, Fractal Image Compression -- Theoy and Application*, Springer Verlag, New York, 1994.

[34] F. Davoine, E. Bertin, J. M. Chassery, *From Rigidity to Adaptive Tessellation for Fractal Image Compression: Comparative Studies, proc. of IEEE 8th Workshop on Image and Mutidimensional Signd processing* 1993.

[35] F. Daroine, E. Bertin, J. M. Chassery, *An Adaptive Partition for Fractal Image Coding, NATO ASI Couf. Fractal Imag Encoding and Analysis*, 1995.

[36] L. Thomas, F. Deravi, *Region Based Fraetal Image Compression Using Heuristic Search, IEEE Trans, on Image processing* Nol. 4. No. 6 1995.

[37] D. Saupe, M. Ruhl, *Evolutionaly Fractal Image Compression*, Proc. ICIP—96 IEEE International Conference on Image Processing, 1996.

[38] M. F. Barnsley, *Lecture Notes on Iterated Function System, Proceedings of Symposia in Applired Mathematics* Vol. 39, 1989.

[39] Emmanuel Rousens, *Sequence Coding Based on The Fractal Theoy of Iterated Transformations Systems SPIE* Vol. 2094.

[40] I. Daubechies , *Ten lectures on Wavelets, Society for Industrial and Applied Mathematics*, 1992.

[41] S. E. Mallat, *A Theoy for Multiresolution Signal Decomposition The Wavelet Represantation, IEEE Trans on Pattern Analysis aud Machine Intelligencl* Vol. 11, No. 7 1989.

[42] E. H. Adelson and E. Simoncelli, *Orthogonal Pyramid Tramsform for Image Coding, Visual Communications and Image Processing* I *Proc.*

SPIE Vol. 645, 1987.

[43] M. Antonini, M. Barland, P. Mathieu and I. Daubechies, *Image Coding Using Wavelet Transform*, *IEEE Trans on Image processing* Vol. 1, No. 2, 1992.

[44] G. Davis, *Self-quantized Wavelet Subtrees: A Wavelet -- based Theoy for Fractal Image Compression*, *proceedings DCC'95 Data Compression couference*, 1995.

[45] G. Davis, *Adaptive. Self Quantization of Wavelet Subtrees, A Wavelet for Fractal Image Compression*, SPIE Comferend on Mathematical Imaging, 1995.

[46] W. Ahmed, T. Natarajan and K. R. Rao, *Discrete Cosine Transform*, *IEEE Trans comput*. Vol. c23, 1974.

[47] Y. Zhao and B. Yuan, *Image Compression Using Fractal and Discrete Cosine Transform*, *Electronics letters*. Vol. 30 No. 6, 1994.

[48] Yipeng Shi, Wei Gu and Shouji Chen, Liming Zhang, *Some New Methods to Fractals Image Compression*, *Communications In Nonlinear Science & Numerical Simulution*, Vol 2, 1997.